

Tracking and forecasting community responses to climate perturbations in the California Current Ecosystem

--Manuscript Draft--

Manuscript Number:	PCLM-D-21-00040R1
Article Type:	Research Article
Full Title:	Tracking and forecasting community responses to climate perturbations in the California Current Ecosystem
Short Title:	Tracking and forecasting community state
Corresponding Author:	Mary Hunsicker NOAA Fisheries Newport, OR UNITED STATES
Order of Authors:	Mary E. Hunsicker Eric J. Ward Michael A. Litzow Sean C. Anderson Chris J. Harvey John C. Field Jin Gao Michael G. Jacox Sharon Melin Andrew R. Thompson Pete Warzybok
Keywords:	Climate variability; climate change; marine heat wave; extreme events; regime shift; ecosystem state; novel state; forecasting; multivariate time series analysis; Dynamic Factor Analysis; north Pacific Ocean; California Current; seabirds; marine mammals; fish
Abstract:	Ocean ecosystems are vulnerable to climate-driven perturbations, which are increasing in frequency and can have profound effects on marine social-ecological systems. Thus, there is an urgency to develop tools that can detect the response of ecosystem components to these perturbations as early as possible. We used Bayesian Dynamic Factor Analysis (DFA) to develop a community state indicator for the California Current Ecosystem (CCE) to track the system's response to climate perturbations, and to forecast future changes in community state. Our key objectives were to (1) summarize environmental and biological variability in the southern and central regions of the CCE during a recent and unprecedented marine heatwave in the northeast Pacific Ocean (2014–2016) and compare these patterns to past variability, (2) examine whether there is evidence of a shift in the community to a new state in response to the heatwave, (3) identify relationships between community variability and climate variables; and (4) test our ability to create one-year ahead forecasts of individual species responses and the broader community response based on ocean conditions. Our analysis detected a clear community response to the marine heatwave, although it did not exceed normal variability over the past six decades (1951–2017), and we did not find evidence of a shift to a new community state. We found that nitrate flux through the base of the mixed layer exhibited the strongest relationship with species and community-level responses. Furthermore, we demonstrated skill in creating forecasts of species responses and community state based on estimates of nitrate flux. Our indicator and forecasts of community state show promise as tools for informing ecosystem-based and climate-ready fisheries management in the CCE. Our modeling framework is also widely applicable to other ecosystems where scientists and managers are faced with the challenge of managing and protecting living marine resources in a rapidly changing

	climate.
Opposed Reviewers:	
Additional Information:	
Question	Response
Financial Disclosure Enter a financial disclosure statement that describes the sources of funding for the work included in this submission and the role the funder(s) played. This includes grants and any commercial funding of the work or authors. This statement will be typeset if the manuscript is accepted for publication. <i>Please review the submission guidelines and the instructions link below for detailed requirements and guidance.</i>	Funding for this project came from NOAA's Fisheries and the Environment (FATE) program (project 16-01) awarded to M.E.H, E.J.W., M.A.L. and C.J.H. and NOAA's California Current Integrated Ecosystem Assessment program (C.J.H.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Competing Interests On behalf of all authors, disclose any competing interests that could be perceived to bias this work. This statement will be typeset if the manuscript is accepted for publication. <i>Please review the instructions link below and PLOS Climate's competing interests policy to determine what information must be disclosed at submission.</i>	The authors do not have any competing interests (financial or non-financial) with respect to the work presented in this research article.
Data Availability Before publication, Authors are required to make fully available and without restriction all data underlying their findings. Please see our PLOS Data Policy page for detailed information on this policy.	All time series collected through the CalCOFI and RREAS surveys are available through the ERDDAP data server (https://coastwatch.pfeg.noaa.gov/erddap/index.html). The seabird and sea lion time series or points of contact are available through the California Current Integrated Ecosystem Assessment (CCIEA) web dashboard (https://www.integratedecosystemassessment.noaa.gov/regions/california-current).

A Data Availability Statement, detailing where the data can be accessed, is required at first submission. Insert your Data Availability Statement in the box below.

Please see the [data reporting](#) section of our submission guidelines for instructions on what you need to include in your Data Availability Statement.

This statement will be typeset if the manuscript is accepted for publication.

PLOS allows rare exemptions to address legal and ethical concerns. If you have legal or ethical restrictions, please use the box below to detail these in full sentences for the Journal team to consider.

December 9, 2021

Dear Dr. Wu,

Thank you for giving us the opportunity to revise our manuscript for PLOS Climate. My co-authors and I have addressed all of the reviewers' comments, as well as the notes from the editorial team regarding the journal's requirements.

We think the reviewers' suggestions were constructive for clarifying our study approach, methodology and results, and we accepted most of them. We note, however, that we were less inclined to remove/re-organize text pertaining to the forecasting component of the paper, as suggested by reviewer #2. The forecasting component is a novel development of our modeling approach and therefore we think it's appropriate to include most of the existing text. We believe this information is valuable for readers who are interested in applying a similar approach to their own study systems.

Thank you for considering our manuscript for publication. We look forward to hearing from you.

Sincerely, on behalf of all co-authors,



Mary E. Hunsicker

1 Short Title: Tracking and forecasting community state

2

3 **Tracking and forecasting community responses to climate perturbations in the California**

4 **Current Ecosystem**

5

6 Mary E. Hunsicker^{1,†,*}, Eric J. Ward^{2,†}, Michael A. Litzow³, Sean C. Anderson⁴, Chris J.
7 Harvey², John C. Field⁵, Jin Gao⁶, Michael G. Jacox⁷, Sharon Melin⁸, Andrew R. Thompson⁹,
8 Pete Warzybok¹⁰

9

10

11 ¹Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National
12 Oceanic and Atmospheric Administration, 2032 SE OSU Drive, Newport, OR 97365 US

13 ²Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service,
14 National Oceanic and Atmospheric Administration, Seattle, WA 98112 USA

15 ³Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric
16 Administration, Kodiak, AK 99615 USA

17 ⁴Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC V9T 6N7 Canada

18 ⁵Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service,
19 National Oceanic and Atmospheric Administration, Santa Cruz, CA USA

20 ⁶Centre for Fisheries Ecosystems Research, Memorial University of Newfoundland, St. John's, NL A1C
21 5R3 Canada

22 ⁷Environmental Research Division, Southwest Fisheries Science Center, National Marine Fisheries
23 Service, National Oceanic and Atmospheric Administration, Monterey, CA 93940 USA / Physical
24 Sciences Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305 USA.

25 ⁸Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service,
26 National Oceanic and Atmospheric Administration, Seattle, WA 98115 US

27 ⁹Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and
28 Atmospheric Administration, La Jolla, CA USA

29 ¹⁰Point Blue Conservation Science, Petaluma, CA USA

30

31 *Corresponding author

32 Email: mary.hunsicker@noaa.gov

33

34 [†] These authors contributed equally to this work

35

36 **Abstract**

37 Ocean ecosystems are vulnerable to climate-driven perturbations, which are increasing in
38 frequency and can have profound effects on marine social-ecological systems. Thus, there is an
39 urgency to develop tools that can detect the response of ecosystem components to these
40 perturbations as early as possible. We used Bayesian Dynamic Factor Analysis (DFA) to develop
41 a community state indicator for the California Current Ecosystem (CCE) to track the system's
42 response to climate perturbations, and to forecast future changes in community state. Our key
43 objectives were to (1) summarize environmental and biological variability in the southern and
44 central regions of the CCE during a recent and unprecedented marine heatwave in the northeast
45 Pacific Ocean (2014–2016) and compare these patterns to past variability, (2) examine whether
46 there is evidence of a shift in the community to a new state in response to the heatwave, (3)
47 identify relationships between community variability and climate variables; and (4) test our
48 ability to create one-year ahead forecasts of individual species responses and the broader
49 community response based on ocean conditions. Our analysis detected a clear community
50 response to the marine heatwave, although it did not exceed normal variability over the past six
51 decades (1951–2017), and we did not find evidence of a shift to a new community state. We
52 found that nitrate flux through the base of the mixed layer exhibited the strongest relationship
53 with species and community-level responses. Furthermore, we demonstrated skill in creating
54 forecasts of species responses and community state based on estimates of nitrate flux. Our
55 indicator and forecasts of community state show promise as tools for informing ecosystem-based
56 and climate-ready fisheries management in the CCE. Our modeling framework is also widely
57 applicable to other ecosystems where scientists and managers are faced with the challenge of
58 managing and protecting living marine resources in a rapidly changing climate.

59 **Introduction**

60 Climate perturbations can have strong impacts on ocean ecosystems that in turn affect social and
61 economic components of human communities. These effects may be exacerbated when changes
62 in ocean conditions are more extreme, such as during marine heatwaves (prolonged events of
63 anomalously warm ocean waters). The increasing attention on these extreme events and their
64 impacts (e.g., Hobday et al. 2018, Sen Gupta et al. 2020) has invigorated a push for tools that can
65 track and detect as early as possible the response of marine communities to climate-driven
66 perturbations. Early detection, and moreover, near-term forecasts of community shifts could help
67 scientists, managers, and stakeholders better prepare for and respond to the potential
68 consequences of such shifts.

69 Climate-driven shifts in community structure tend to involve rapid change across
70 multiple populations that result in switches between contrasting community assemblages that
71 may then persist for decades. A growing number of studies have documented community
72 reorganizations in response to climate drivers (e.g., Beaugrand et al. 2008, 2015, Möllman and
73 Diekmann 2012, Wernberg et al. 2016, Peabody et al. 2018). One of the best-known examples is
74 the widespread northeast Pacific community reorganization that followed the 1976/1977 shift in
75 the Pacific Decadal Oscillation from a cold to warm regime (Benson and Trites 2002; Hare and
76 Mantua 2000). The abrupt change from a cool to warm ocean regime had dramatic implications
77 on ecosystem functioning and living marine resources (LMRs) throughout the region (Mantua et
78 al. 1997, Anderson and Piatt 1999; Litzow and Ciannelli 2007, Peabody et al. 2018). Since then,
79 northeast Pacific marine ecosystems have experienced several interannual or decadal
80 perturbations that do not appear to have resulted in community-wide shifts of similar magnitude.

81 However, between 2014 and 2016 these ecosystems experienced a marine heatwave that
82 involved the warmest sea surface temperature (SST) and heat content anomalies that had ever
83 been observed over large areas of the North Pacific, with SST anomalies over 6°C (Bond et al.
84 2015; Walsh et al. 2018). It was one of the most extreme heatwaves globally in its combined
85 magnitude, spatial scale, and duration (Hobday et al. 2018, Sen Gupta et al. 2020), and the
86 intense, persistent warming has been attributed to a combination of natural and anthropogenic
87 forcing (Jacox et al. 2018a; Laufkötter et al. 2020). Several studies have documented myriad
88 biological responses to this event. For example, within the California Current Ecosystem (CCE),
89 there were mass strandings of marine mammals (Cavole et al. 2016), increased whale
90 entanglements due to shifting prey sources (Santora et al. 2020), mass mortality events for
91 marine seabirds (Cavole et al. 2016, Jones et al. 2018, Piatt et al. 2020), a record-breaking
92 domoic acid outbreak (McCabe et al. 2016), shifts in pelagic macronekton and micronekton
93 communities and species richness (Santora et al. 2017, Brodeur et al. 2019, Nielsen et al. 2020),
94 irruptions of previously rare fishes and invertebrates throughout the California Current (Sakuma
95 et al. 2016, Morgan et al. 2019, Sanford et al. 2019, Walker et al. 2020), and extraordinarily high
96 recruitment of rockfishes (genus *Sebastodes*; Schroeder et al. 2018, Field et al. 2021) and northern
97 anchovy (*Engraulis mordax*; Thompson et al. 2019). Yet, to date, there have been few
98 quantitative studies of how the marine heatwave impacted the broader CCE community at
99 multiple trophic levels, and therefore the importance of this extreme event for community-wide
100 patterns of variability, and the persistence of the community response, remains largely unknown.

101 Indicators of community or ecosystem state are valuable tools for tracking climate-related
102 changes in ecosystem functioning and evaluating those changes within the context of past
103 climate perturbations (Harvey et al. 2020). Moreover, combining long-term monitoring surveys

104 and data with modeling frameworks that summarize information across taxa and life stages that
105 respond quickly to climate perturbations could provide early detection of an ecosystem shifting
106 into a novel state. Early detection of such shifts would benefit ecosystem-based and climate-
107 ready fisheries management strategies aimed at mitigating possible deleterious ecological and
108 socio-economic outcomes. There is also a pressing need for forecasts of future ecosystem states
109 to support forward-looking management of LMRs (Hobday et al. 2016, Tommasi et al. 2017,
110 Jacox et al. 2020), including assessments of risk. As climate models and forecasts of ocean
111 conditions continue to improve, there are burgeoning opportunities to develop and test methods
112 that could provide near-term forecasts of community state in relation to ocean conditions.

113 A challenge in summarizing ecosystem responses to perturbations is that time series used
114 to characterize the ecosystem often involve tens to hundreds of variables (species or climate
115 indices); there is often some degree of asynchrony among time series (unevenly or irregularly
116 spaced), and further, each is corrupted by the presence of observation errors. Disentangling these
117 sources of error and separating the signal from the noise is statistically challenging.

118 Traditionally, tools such as Principal Components Analysis (PCA) or nonmetric
119 multidimensional scaling have often been used for identifying leading patterns of variability in
120 multivariate datasets (e.g., Koslow et al. 2002, 2013); however, these approaches are ill-suited to
121 the analysis of time series data that are autocorrelated or non-stationary (Planque and Arneberg
122 2018). An alternative approach, Dynamic Factor Analysis (DFA), is better suited for identifying
123 shared trends that can be used as a community state indicator. DFA is specifically designed for
124 time series ordination, and avoids many of the problems associated with other multi-variate
125 approaches (Zuur et al. 2003). When applied to a collection of multivariate time series, inference
126 in DFA models focuses on estimating a smaller number of temporal patterns ('trends') that best

127 capture the variation observed. The observed data are then treated as a mixture of these trends
128 (Ward et al. 2019). Ward et al. (2019) recently developed a Bayesian implementation of DFA
129 that offers added flexibility in model aspects over conventional approaches; examples include
130 allowing for extreme “black swan” events (rare and difficult to predict events; Anderson et al.
131 2017), and trend processes that do not follow a random walk. Output from these Bayesian DFA
132 models can also be used to estimate the probability of extreme events occurring or switches
133 among contrasting system states. In the first application of this new method, Litzow et al.
134 (2020a) examined shared trends of climate and biology time series in the Gulf of Alaska. Their
135 study did not detect evidence for wholesale community reorganization during the recent
136 northeast Pacific marine heatwave; however, their findings indicated potential for new patterns
137 of ecosystem functioning with continued warming of ocean temperatures.

138 The goal of our study is to build on this set of novel statistical tools to develop a model of
139 the CCE state that can both track and forecast ecosystem changes in response to climate
140 perturbations. More specifically, we expand the Bayesian implementation of DFA to test the
141 community response to environmental variables within the modeling framework and to develop
142 near-term forecasts of future community states. Using climate and biological data from the
143 central and southern regions of the CCE, our specific objectives were to: (1) summarize
144 environmental and biological variability during 2014–2016 marine heatwave and compare these
145 patterns to past variability; (2) examine whether there is evidence of departures from previous
146 climate patterns and of switches to a new community state during the heatwave; (3) identify
147 relationships, if any, between community variability and climate variables; and (4) test our
148 ability to create one-year ahead forecasts of species responses and the community state based on
149 environmental information. While the focus of our study is the CCE, the approach applied here is

150 widely applicable to the myriad marine ecosystems worldwide that are vulnerable to a rapidly
151 changing climate.

152

153 **Methods**

154 *Data*

155 In our analysis, we used oceanographic time series from the southern (n=6) and
156 central (n=6) regions of the CCE, derived from a data assimilative configuration of the Regional
157 Ocean Modeling System (ROMS) with 0.1° (~ 10 km) horizontal resolution and 42 terrain-
158 following vertical levels (Neveu et al. 2016; oceanmodeling.ucsc.edu). From the ROMS output,
159 we generated monthly time series covering 1980-2018 for a suite of variables including sea
160 surface temperature (SST), sea surface height (SSH), isothermal layer depth (ILD), Brunt-
161 Väisälä frequency (BV), a coastal upwelling transport index (CUTI), and a biologically effective
162 upwelling transport index (BEUTI). The ILD is similar to mixed layer depth and defines the
163 depth where temperature deviates by 0.5°C from the surface value. BV is a measure of water
164 column stratification, averaged over the upper 200 m of the water column. CUTI and BEUTI are
165 upwelling indices that quantify vertical transport and nitrate flux through the base of the mixed
166 layer, respectively (Jacox et al. 2018b). The data were annually averaged (July-June) from the
167 coast to 100 km offshore, with the exception of CUTI and BEUTI, which capture coastal
168 upwelling within 75 km of shore. In the alongshore direction, we calculated averages for two
169 regions with a division at Point Conception, California, separating the southern portion of the
170 CCE ($31\text{--}34.5^\circ\text{N}$) from the central region ($34.5\text{--}40.5^\circ\text{N}$, Fig. 1). This is in response to the
171 recognition of Point Conception as a major biogeographic boundary for the California Current
172 System, with differing wind and current patterns north and south of that feature (Checkley and

173 Barth 2009, Gottscho 2016). The annual averages were taken from July to June to capture the
174 influence of the El Niño–Southern Oscillation (ENSO), which peaks in winter and is the
175 dominant mode of interannual variability influencing the California Current (Jacox et al. 2015).
176 We developed models using ROMS output rather than empirical measurements because they
177 provide full spatial and temporal coverage of surface and subsurface conditions, incorporate
178 available observations, and will enable the use of ROMS forecasts to then forecast biological
179 changes in the CCE. This ocean model is constrained by available satellite and *in situ*
180 observations to improve its fidelity to nature and has been validated against independent *in situ*
181 observations (Neveu et al. 2016, Schroeder et al. 2014). Output from this model has been widely
182 used to characterize CCE oceanography, its relation to large scale climate variability, and its
183 influence over the marine ecosystem from phytoplankton to top predators (see Discussion). More
184 details on the oceanographic time series can be found in S1 Table and S1 Figure.

185

186 **Figure 1.** Sampling locations of California Current Ecosystem biology included in the study
187 analyses. Abundance data for pelagic juvenile groundfishes and invertebrates are collected on the
188 Rockfish Recruitment and Ecosystem Assessment Survey (RREAS). Ichthyoplankton data are
189 collected on the California Cooperative Oceanic Fisheries Investigations (CalCOFI) survey.
190 Seabird reproductive success and California sea lion (*Zalophus californianus*) pup time series are
191 collected on Southeast Farallon Island and San Miguel Island, respectively. See S1 Table and S1
192 Figure for detailed information on the individual time series. The base map layer was sourced
193 from [NOAA National Geophysical Data Center \(2009\) ETOPO1 1 Arc-Minute Global Relief](#)
194 [Model](#). NOAA National Centers for Environmental Information (accessed: 19 April 2013,
195 Amante, C & BW Eakins 2009).

196

197 The biology time series included in our analysis were selected based on three criteria: first, the
198 measured variables would be expected to show rapid (0- to 1-year lag) responses to climate
199 variability; second, the time series could be updated with no more than one year lag for
200 processing time to increase the speed at which biological responses to perturbation could be
201 detected; and third, the time series were at least 15 years long. A threshold of 15 years allowed
202 us to include long time series that spanned as many climate perturbations as possible and also
203 have enough biological time series to develop an informative indicator of community state. In
204 addition, 15 years is a threshold that has been previously used to define "long oceanographic
205 time series" in the California Current (McClatchie et al. 2014). The biology time series that met
206 our selection criteria (n=38) included ichthyoplankton, pelagic young-of-the-year (juvenile fish),
207 squid, and krill abundance; seabird productivity; and California sea lion pup body condition
208 metrics (Fig. 1, S1 Table). These 38 time series were collected from four disparate ocean
209 surveys, and span between 22 and 68 years. Datasets collected from surveys that included spatial
210 attributes (e.g., ichthyoplankton and pelagic juvenile fish surveys) were first standardized using
211 Generalized Additive Models to create a univariate time series for each species. While these
212 datasets generally include spatial random sampling, the index standardization accounts for
213 uneven distributions of effort (in space or time). Details on the standardization of individual
214 datasets are included in S1 Appendix. In addition, the biology data were normalized with log
215 transformations where appropriate (all zeros were changed to NAs). For example, if the time
216 series data were assumed to be lognormally distributed (e.g., weight/count data) or the
217 coefficient of variation was > 1, the data were log transformed. All of the time series from an
218 individual dataset (survey) were treated the same, i.e., logged or not. More details on the biology

219 time series used in this study and the associated data sources and log transformations are
220 summarized in S1 Table and S1 Figure.

221

222 *Modeling*

223 We describe the methods in detail below, but in summary our work flow was to (1) apply
224 Bayesian DFA to climate and biology datasets separately and use model selection tools to
225 identify the best supported model and number of shared trends, (2) apply ‘black swan’ and
226 regime detection methods to detect extreme events and alternating community states,
227 respectively, (3) identify whether the CCE community state was strongly correlated with the
228 climate time series (compare performance of the biology models with/without environmental
229 covariates), and (4) evaluate our skill at making predictions of community state and individual
230 species variables. These four steps map on to the four study objectives outlined in the
231 introduction.

232

233 *Dynamic Factor Analysis*

234 We used a Bayesian version of Dynamic Factor Analysis (DFA, Zuur et al. 2003, Ward et
235 al. 2019) using the software Stan and R (R Core Team 2018) as implemented in the ‘bayesdfa’
236 package (Ward et al. 2020). DFA is a multivariate statistical tool somewhat analogous to
237 Principal Components Analyses, but for time-series data (Holmes et al. 2018, <https://cran.r-project.org/web/packages/MARSS/vignettes/UserGuide.pdf>). For a collection of time series, the
238 number of estimated ‘trends’ is specified *a priori*, and DFA estimates these latent trends as
239 independent random walks. In mathematical form, this is expressed as
240

$$x_t = x_{t-1} + w_{t-1},$$

242 where \mathbf{x}_t represents the value of latent (unobserved) trends at time t , and the process error
243 deviations \mathbf{w}_{t-1} are generally assumed to be white noise having arisen from a multivariate
244 normal distribution (with an identity covariance matrix for identifiability). The latent trends are
245 mapped to the observed data through an estimated loadings matrix \mathbf{Z} and residual error \mathbf{e}_t ,

246
$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{b} \cdot \mathbf{d}_t + \mathbf{e}_t,$$

247 where \mathbf{y}_t is the vector of observed states at time t , and the residual error terms \mathbf{e}_t are assumed to
248 be drawn from a univariate or multivariate normal distribution. Though the covariance matrix of
249 \mathbf{w}_t is generally fixed (Zuur et al. 2003), the covariance matrix of \mathbf{e}_t can be structured; variances
250 may be shared or not across time series, and off diagonal elements may be estimated. The
251 parameter vector \mathbf{b} represents optional estimated coefficients relating covariates \mathbf{d}_t to the
252 observed response. In the context of our DFA modeling, we included climate variables as \mathbf{d}_t in
253 models where the biological observations were used as the response \mathbf{y}_t .

254 Because we implemented the DFA model in a Bayesian setting, we were able to extend
255 this model to include additional features. First, to include extreme events, we relaxed the
256 assumption about process errors \mathbf{w}_t being drawn from a normal distribution and used a
257 multivariate Student-t distribution (MVT) instead (Anderson and Ward 2019). We also modified
258 the process equation to consider an optional vector of AR(1) coefficients $\boldsymbol{\phi}$ on the latent trends.
259 $\mathbf{x}_t = \boldsymbol{\phi}\mathbf{x}_{t-1} + \mathbf{w}_{t-1}$ (Ward et al. 2019). A final modification of the conventional DFA model is
260 that for some models, process variances can be estimated rather than fixed at 1 (maximum
261 likelihood approaches generally use this constraint for identifiability). As implemented in Stan
262 (Stan Development Team 2016, Hoffman and Gelman 2014, Carpenter et al. 2017), we
263 conducted estimation with three chains, with a warm-up of 2000 samples, followed by 2000
264 iterations. We used the split-chain potential scale reduction factor (Gelman and Rubin 1992,

265 Gelman et al. 2013) to assess convergence ($Rhat < 1.05$). Code to replicate these analyses is
266 deployed as an R package on CRAN ('bayesdfa', Ward et al. 2020) and our public Github
267 repository, <https://github.com/fate-ewi/bayesdfa>.

268

269 *Models structure optimization*

270 We ran the DFA on climate datasets (1981–2017) and biological datasets (1951–2017)
271 across the southern and central regions of the California Current combined. Running the analysis
272 at this spatial scale allowed us to capture the broader community response to climate
273 perturbations, compared to running models on each multivariate dataset independently (e.g., time
274 series from a single survey). There are a number of ways to evaluate predictive accuracy of these
275 models. The commonly used Leave-One-Out Cross-Validation (LOO-CV), for example holds
276 each observation out in turn and predictions are made from the remaining data. As our focus was
277 on the temporal nature of the data and forecasting component, we implemented a variant of k-
278 fold cross validation and treated individual years as unique ‘folds’. Because our objectives
279 involved evaluating these models for future predictions, we implemented the Leave-Future-Out
280 Cross Validation Information Criterion (LFO-CV, Bürkner et al. 2020). We used this approach to
281 identify data support for (1) the number of latent DFA trends ($n = 1-3$), (2) first-order
282 autoregressive AR(1) coefficients on the trends (ϕ estimated with a $Normal(0,1)$ prior), (3)
283 Student-t deviations (i.e., evidence of extreme events, using a prior on the MVT degrees of
284 freedom parameter, ν , of $\nu \sim Gamma(2, 0.1)$), and (4) a fixed versus estimated trend variance
285 (using a prior on the standard deviation, σ_w , of $\sigma_w \sim Normal(0,1)$).

286 In addition, we used LFO-CV to identify the most appropriate error structure for the
287 climate dataset—specifically whether the times series had equal (shared) or unequal (unique)

288 observation errors. For the biology models, we assumed the observation errors were unique by
289 dataset, and our estimates of survey variance supported this assumption.

290 For each model formulation, we applied the LFO-CV method by first fitting the model to
291 all years of data prior to year T (i.e., training data, years 1, 2, ..., $(T-1)$) and then using the fitted
292 model to predict the trend value in year T (i.e., test data). We repeated this process for 10 years,
293 starting with 2017 as year T and working back to 2008, and then calculated the expected log
294 predictive density (ELPD) across those time steps. The climate and biology models with the
295 highest ELPD were deemed the best supported models. The LFO-CV is a preferred method for
296 evaluating future predictive performance of Bayesian models because it properly accounts for
297 time series structure, and unlike other Bayesian cross-validation methods, does not produce
298 overly optimistic estimates (Bürkner et al. 2020).

299

300 *Detection of extreme events and regime shifts*

301 After identifying the best-supported DFA model for the climate and biological datasets,
302 we conducted a post-hoc examination of outlier detection and regime shifts. For outlier detection
303 of black swan events, we implemented a method similar to that described in Anderson et al.
304 (2017) and applied it to the climate and biology time series. This approach relies on first
305 differencing the posterior trend mean estimates of the climate and biology trends, $x_t - x_{t-1}$ and
306 then applies a normal density function to identify year-over-year changes that were unlikely to
307 have arisen from a normal distribution (given the process variance). Probabilities can then be
308 assigned to the deviations in each year (e.g., ‘there is a 1:1000 chance of observing a deviation
309 similar to that estimated in year t ’). As described in Ward et al. (2019), the presence of regimes
310 can also be estimated by applying hidden Markov models (HMM) to the estimated state indices

311 from a DFA. We evaluated support for regimes and alternate states by using the posterior trend
312 estimates from each model as input. The Bayesian Leave-One-Out Cross Validation information
313 criterion (LOO-CV, Vehtari et al. 2017) was used to identify the data support for the number of
314 regimes ($n = 1–3$). The model with the lowest LOO-CV value is deemed the best model.

315

316 *Climate-biology relationships and forecasts of community state*

317 While a wide variety of multivariate or univariate time series methods could be applied to our
318 observed time series to generate forecasts, our objectives were to develop simultaneous estimates
319 of both the community state (i.e., the DFA trend value) and the raw time series (i.e., individual
320 time series summarized by the biology DFA model). We evaluated the ability of our DFA
321 models to generate short-term (one year lead-time) forecasts of community state by first
322 evaluating whether the performance of the biology DFA model was improved when climate time
323 series were included as covariates in the model. If climate time series were found to better
324 explain the variability in the biology time series, these relationships could potentially be used to
325 forecast community trends. For our analysis, we ran the DFA on a subset of the biology data
326 overlapping in time with the climate dataset, i.e., 1981–2017, to make out-of-sample predictions.
327 We used the same LFO-CV procedure described above, with the same forecast period (2009–
328 2017) to compare the biology models with and without a single climate covariate (see S2 Table
329 for all model formulations). In this case, the model used biological and climate data from all
330 preceding years and climate data from the year to be forecast. The six climate covariates from
331 the southern region and the central region of the CCE (12 total) were tested in this analysis. Once
332 the best-supported biology-covariate model was identified, we used that model to make
333 predictions of the community state (i.e., DFA trend value) in 2018 using climate data from that

334 same year and the raw time series of the individual species (i.e., the biology time series
335 summarized by the DFA model). We evaluated forecast skill based on the prediction errors of
336 individual species time series and by comparing the forecasts for 2008–2018 to the 2008–2018
337 trend values estimated from the biology-covariate model that only included data prior to the
338 forecast year.

339

340 **Results**

341 *Climate and biology trends*

342 The model with the highest predictive accuracy (ELPD) of the climate state in the southern and
343 central regions of the CCE was a one-trend DFA model (Model 1 in Table 1, Fig. 2a). This
344 model included unique observation variances across the six time series, support for heavy-tailed
345 deviations of the latent trend, an AR(1) coefficient on the trend (S1 Fig.), and an estimated trend
346 variance. Overall, the trend captured a well-documented cooling period in the CCE between
347 1980 and 2010 (e.g., Seo et al. 2012), as well as strong El Niño events (e.g., 1982–1983, 1997–
348 1998, 2015–2016) and the 2014–2016 marine heatwave. The trends and loadings indicate that
349 these events were generally associated with weaker upwelling, reduced mixed layer depth, low
350 nutrient flux, and warm, stratified waters (Fig. 2a, b). All but one of the climate time series
351 (central ILD) were strongly associated with the single trend, i.e., at least 90% of the loading
352 posterior distributions associated with each time series were above or below zero (Fig. 2b). The
353 SST, SSH, and BV frequency (water column stratification) time series from the southern and
354 central regions of the CCE loaded positively on this trend (Fig. 2b). The BEUTI and CUTI time
355 series from both regions of the CCE and the ILD time series from the central region loaded
356 negatively on the trend (Fig. 2b).

357

358 **Table 1.** Summary information for climate and biology Bayesian DFA models, including
 359 whether process error was estimated, observation error variances (unequal or equal among time
 360 series, or unique to each survey), the number of model trends, expected log pointwise predictive
 361 densities (ELPD), and standard error of ELPD. Bold text highlights the models that show best
 362 support or highest predictive accuracy for the climate and biology data for the southern and
 363 central California Current ecosystem (i.e., highest ELPD). All climate and biology models
 364 include an AR(1) process and Student-t deviations.

Time series	Model	Process sigma	Variance index	Trends	ELPD	SE ELPD
Climate	1	Yes	unequal	1	-10551.89	759.44
	2	No	unequal	1	-10682.39	712.96
	3	Yes	equal	1	-16793.59	1732.54
	4	No	equal	1	-16881.03	1824.42
	5	No	unequal	2	-17441.01	1655.23
	6	Yes	unequal	2	-17818.86	1813.65
	7	Yes	unequal	3	-21571.65	1674.77
	8	No	unequal	3	-22882.99	1695.12
	9	No	equal	2	-23927.86	2694.72
	10	Yes	equal	2	-24536.19	2953.38
	11	Yes	equal	3	-38895.66	4024.93
	12	No	equal	3	-38920.87	4762.15
Biology	13	No	survey	1	-2003.32	95.76
	14	Yes	survey	1	-2003.96	88.88
	15	No	survey	2	-2155.52	76.05
	16	Yes	survey	2	-2198.79	78.86
	17	No	survey	3	-2334.55	118.31
	18	Yes	survey	3	-2335.10	112.70

365

366

367 **Figure 2.** Climate variability in southern and central California Current ecosystem: a) shared
 368 trend with 95% credible intervals (1981–2017), b) posterior distributions for loadings on all of
 369 the individual time series. Loadings with darker shading indicate time series loading most
 370 strongly on the climate trend. SST, sea surface temperature; SSH, sea surface height; ILD,
 371 isothermal layer depth; BV, Brunt-Väisälä frequency (stratification); CUTI, Coastal Upwelling

372 Transport Index; BEUTI, Biologically Effective Upwelling Transport Index. See S1 Table and
373 S1 Figure for climate times series details.

374

375 The climate state during the marine heatwave, as indicated by the DFA trend, was within
376 the bounds of previous observations. While there was support in the best model for heavy-tailed
377 deviations in the climate trend (i.e., Student-t deviations S2 Fig.), our post-hoc examination of
378 outliers detected a single extreme event in the climate state in mid-1998 to mid-1999
379 (threshold=0.001), when there was a shift from strong El Niño (1997–98) to strong La Niña
380 (1998–1999) conditions, and not around the time of the heatwave. Application of the Bayesian
381 HMM to the climate trend most supported the presence of two hidden states, reflecting the
382 probability of being in a state associated with warmer conditions versus one with cooler
383 conditions (LOO-CV: one-state = 129.1, two-state = 9.4, three-state = 27.2, Fig. 3). The LOO-
384 CV did not provide support for a shift to a third novel climate state in the southern and central
385 regions of the CCE during the marine heatwave, however there is a shift back to the previously
386 observed warm state during the marine heatwave.

387

388 **Figure 3.** Results of Hidden Markov Models (HMM) showing state probability for latent trends
389 in the climate (top) and biology (bottom) data sets. The best model for both data sets invoked
390 two states, and the median probability (and 95% credible intervals) of being in one state versus
391 the other is shown. The top figure reflects the probability of being in the state associated with
392 warmer conditions versus one with cold conditions, and bottom figure indicates that ecosystem
393 did not shift into a new state following the marine heatwave.

394

395 The best model for community variability among our biological time series was also a
396 one-trend model (Model 13 in Table 1, Fig. 4a). The model formulation was similar to the best
397 climate model, except the observation variances were unique by dataset (survey) and not
398 individual time series. We note that the top two models (Model 13 and 14) showed similar
399 predictive accuracy ($\Delta \text{ELPD} < 1$) and only differed with respect to whether the process variance
400 was fixed at 1 or estimated. Here we only show results for the model with a fixed process
401 variance. The biology showed strong coherence in community signal; a majority of the time
402 series (31 of 38) loaded strongly (probability > 0.9) on the single trend and most of them
403 demonstrated loadings in the same direction (Fig. 4b). The magnitude and direction of the
404 estimated loadings were consistent with the observed high relative abundance of most juvenile
405 groundfishes (rockfish, flatfish), squid, krill, and some ichthyoplankton species during the
406 marine heatwave, and suggest that the reproductive success of some seabird species was higher
407 around the time of the heatwave as well. The few time series loading in the other direction on the
408 trend indicated a reduction in sea lion pup growth rate and lower abundances of juvenile/adult
409 Pacific sardine *Sardinops sagax* and some ichthyoplankton species (e.g., larval northern anchovy
410 and Pacific hake *Merluccius productus*) associated with the heatwave.

411
412 **Figure 4.** Community variability in the southern California Current ecosystem: a) shared trend
413 with 95% credible intervals (1951–2018: marine heatwave occurred 2014–2016), b) posterior
414 distributions for loadings on individual time series (only time series with $\geq 90\%$ of the loading
415 distributions above or below zero are shown). Loadings with darker shading indicate time series
416 loading most strongly on the biology trend. Cal. = California, Juv. = juvenile fish stage,

417 Juv./adult = juvenile and adult fish stages combined, all other fish are larval fish. See S1 Table
418 and S1 Figure for times series details.

419

420 The estimated trend from this biology DFA model demonstrates a potential shift in community
421 state in the mid-1960s, although there is considerable uncertainty around the trend during this
422 early portion of the time series, likely due to the limited number of observations
423 (ichthyoplankton only) pre-dating the 1970s (Fig. 4a, S1 Fig.). The community state appears to
424 be relatively stable from the late 1970s through the early 2000s, and the trend reached a peak
425 around 2013–2015. Evidence of a community shift early in the time series is supported by our
426 regime detection analysis, which demonstrated that a two-state model best described the latent
427 trend (LOO-CV: one-state = 216.4, two-state = 11.8, three-state = 41.8, Fig. 3). This shift
428 coincides with a strong increase in the abundance of a few species during that period, including
429 eared blacksmelt (*Lipolagus ochotensis*), slender blacksmelt (*Bathylagus pacificus*), northern
430 lampfish (*Stenobrachius leucopsarus*), which are cool water associated mesopelagic species, as
431 well as a rise in northern anchovy (*Engraulis mordax*) abundance prior to the shift (S1 Fig.). Our
432 analysis does not document a shift to a novel community state in response to the recent marine
433 heatwave.

434 While this model provided slight support for heavy-tailed Student-t deviations in the
435 latent trend (S3 Fig.), we did not detect any black swan events in the community state. We note
436 that the community response to two strong El Niño events (1982–1983 and 1997–1998) and to
437 unusually low productivity conditions (2005) in the central CCE appear similar in magnitude and
438 duration to the response to the 2014–2016 marine heatwave, although the directions of the
439 responses were opposite (Fig. 4a). Our regime detection analysis also captured the change in the

440 central CCE community in the mid to late 2000s (Fig. 3), which may be associated with the large
441 changes in the reproductive success of multiple seabirds (e.g., Cassin's auklet
442 *Ptychoramphus aleuticus*, common murre *Uria aalge*, Brandt's cormorant *Urile penicillatus*) and
443 in sea lion pup births around that time (S1 Fig). These taxa may have been impacted by changes
444 in the abundance or availability of important prey items resulting from unproductive ocean
445 conditions in the central CCE in 2005 and the below normal SSTs associated with the 2007-2008
446 La Niña Event (McClatchie et al. 2008, 2009, Bjorkstedt et al. 2010).

447

448 *Forecast of community state*

449 In comparing models of the biological response with and without climate covariates, we found
450 that several biology models with climate predictors outperformed the biology models that did not
451 include covariates (S2 Table). The climate covariate resulting in the best future predictions of
452 community state was BEUTI (central region), followed by CUTI (central region) (Table 2, see
453 S2 Table for all models). The coefficients linking BEUTI to observed time series indicate strong,
454 positive relationships between nitrate flux and the reproductive success of seabirds and the
455 abundance of krill in the central California Current (Fig. 5). They also indicate strong, negative
456 relationships between nitrate flux and the abundance of juvenile/adult Pacific sardine and larval
457 northern anchovy (Fig. 5). The remaining biology-BEUTI relationships were moderate (e.g.,
458 ichthyoplankton, market squid *Doryteuthis opalescens*) to weak (e.g., rockfish spp., Fig. 5). The
459 biology-CUTI model was similar to the biology-BEUTI model with respect to model structure
460 and estimated species loadings. The estimated coefficients in the CUTI model (S4 Fig.) also
461 show a similar pattern to those in the biology-BEUTI model. The remaining covariate models
462 only showed weak climate-biology relationships (e.g., S5-S6 Fig.).

463

464 **Table 2.** Summary information for the top biology-covariate Bayesian DFA models for each
 465 covariate and the top two biology only models (years 1981-2017). The table indicates whether
 466 process error was estimated ('Yes') or fixed ('No'), the number of model trends, expected log
 467 pointwise predictive densities (ELPD), standard error of ELPD, the environmental covariate
 468 included in the model, and the region in the California Current over which the covariate was
 469 aggregated. All models had an AR(1) coefficient on the trend and Student-t deviations. Model 1
 470 was deemed the best model based on its highest predictive accuracy (highest ELPD value)
 471 compared to all other models. BEUTI = Biologically Effective Upwelling Transport Index; BV =
 472 Brunt-Väisälä frequency; CUTI = Coastal Upwelling Transport Index; ILD = Isothermal Layer
 473 Depth; SST = Sea Surface Temperature; SSH = Sea Surface Height. See S2 table for the full
 474 suite of model comparisons.

475

Model	Process sigma	Trends	ELPD	SE ELPD	Covariate	Region
1	No	1	-1878.71	71.52	BEUTI	central
2	No	1	-1905.27	81.98	CUTI	central
3	No	1	-1914.62	83.22	ILD	south
4	No	1	-1927.33	88.26	SST	south
5	No	1	-1928.03	88.39	BV	south
6	No	1	-1946.75	86.33	SSH	south
7	No	1	-1951.44	73.59	None	-
8	Yes	1	-2038.71	91.44	None	-

476

477

478 **Figure 5.** A summary of the effect of the Biologically Effective Upwelling Transport Index
 479 (BEUTI), a measure of nitrate flux through the base of the mixed layer, on the single species
 480 parameters. Cal. = California, Juv. = juvenile fish stage, Juv./adult = juvenile and adult fish

481 stages combined, all other fish are larval fish. Blue error bars reflect 95% credible intervals.

482 Figures S4-S6 show effects of other environment covariates on the biological variables.

483

484 Given that the biology-BEUTI model was the best supported model over the null model
485 (a model without covariates), we were interested in evaluating the ability of this covariate model
486 to forecast the community state. Comparisons between the community state (out-of-sample
487 estimates) and the community state forecasts indicate that we had skill in forecasting community
488 state one year in advance (Fig. 6, S7 Fig., S8 Fig.). Forecasts of the community trend values for
489 ten additional years (2008–2017, Fig. 7) also indicate that we had some skill for many of the
490 years tested. There are wide confidence intervals around the forecasts; however, given our
491 methodology we can expect that the credible intervals around the trend forecast will be larger
492 than the historical credible intervals (Fig. 6). Forecasts have more uncertainty than historical
493 values because the variance of a random walk increases linearly with time (Holmes 2004; Ward
494 et al. 2014). Furthermore, our credible intervals are increased because we are additionally (1)
495 propagating full parameter uncertainty across the MCMC draws projecting it, and (2) using a
496 Student-t distribution, which has heavy tails and therefore makes the uncertainty intervals wider
497 than if we used normal distribution.

498

499 **Figure 6.** Community variability and forecast of the community state in the southern California
500 Current. The shared biology trend (blue line, with 95% credible intervals) derived from biology-
501 BEUTI model fit to subset of data (1981–2018) is shown along with the trend forecast for 2018
502 (circle, with 95% credible intervals). See S7 Fig. for model loadings.

503

504 **Figure 7.** Forecasts and model estimates of the ‘true’ community state in the southern and
505 central California Current in years 2008—2018 (circle, with 95% credible intervals).

506

507 Overall, the model forecast skill of individual species parameters was moderate to high
508 for half of the species included in the biology-BEUTI model (S3 Table, S8 Fig.). It is important
509 to emphasize that the source of variability in predictions for each of the original time series is a
510 mixture of the magnitude and uncertainty around the trends and loadings (\mathbf{x}_t , \mathbf{Z}), and the
511 magnitude and uncertainty in the estimated covariates (\mathbf{b}). On the one hand, the time series
512 associated with the highest predictive skill (i.e., lowest prediction errors) included seabird
513 reproductive success (common murre, Cassin’s auklet) and the abundance of juvenile Pacific
514 sanddab *Citharichthys sordidus*, juvenile halfbanded rockfish *Sebastodes semicinctus*, market
515 squid, and several ichthyoplankton species (S3 Table, S8 Fig.). On the other hand, forecast skill
516 was lowest (i.e., highest prediction errors) for the abundance of some juvenile rockfishes
517 (chilipepper *Sebastodes goodei* and widow rockfish *Sebastodes entomelas*) and larval fishes (northern
518 anchovy, mesopelagics), which is likely attributed to a lag or mismatch in the timing of the
519 climate-biology relationships. These patterns in forecast skill are similar to those based on the
520 biology-CUTI, -SST, and -ILD models (S3 Table). Lastly, the uncertainty around model
521 predictions of species parameters appears to be driven more by the precision of the model
522 coefficients than by the loadings on the community trend (e.g., S9 Fig.).

523

524 **Discussion**

525 We applied a novel set of statistical tools to data from the southern and central regions of the
526 CCE to document the community response to climate perturbations over the past six decades and

527 to create near-term forecasts of community state. Our analysis detected a community response to
528 the 2014–2016 northeast Pacific marine heatwave; however, it did not exceed normal variability
529 within the study timeframe or result in a shift to a novel community state, based on the biological
530 time series investigated here. We identified relationships between community state and multiple
531 climate variables, with nitrate flux through the base of the mixed layer having the strongest
532 correspondence with individual species time series and the shared trend in community
533 variability. Moreover, we demonstrated skill in creating simultaneous one-year lead time
534 forecasts of species responses and community state.

535

536 *Long-term changes in community state*

537 Many studies and anecdotal accounts have documented unexpected biological responses
538 to the 2014–2016 northeast Pacific marine heatwave. Based on the biological time series
539 included in our analysis, the broader CCE community demonstrated a clear response to the
540 marine heatwave (Fig. 4a). However, our results do not demonstrate a widespread community
541 reorganization beyond the archetypal community structure of this dynamic ecosystem in
542 response to this event. Instead, the mean values for the shared trend in the biology time series, as
543 well as for the shared climate trend, were within the range of previous observations. Many
544 species were present during the marine heatwave that are not typically observed in the CCE.
545 While our analysis could not include these sporadically occurring taxa, due to the large number
546 of zero observations in the historic survey data, the exceptional presences and high abundances
547 of those warm species did not result in a persistent signal among the species included in the
548 DFAs. As additional years of data become available, the DFA models could reveal different
549 outcomes from 2014–2016. However, this is unlikely given that the taxa and life stages used in

550 both studies are known to respond quickly to changes in ocean conditions and given our
551 assumption that the surveys are consistently sampling at the right time and location to fully
552 characterize the short-term response.

553 While our study did not detect a shift in community state in the southern and central CCE
554 during the 2014–2016 heatwave, we did detect a shift in the 1960s. The 1960s shift was likely
555 due to a regime shift previously detected in the southern California ichthyoplankton community
556 (Peabody et al. 2018). The Peabody et al. (2008) study included a much broader suite of
557 ichthyoplankton species than our study which limits our ability to evaluate whether the species
558 driving the shifts are consistent among studies. Previous studies have also documented a shift in
559 response to the 1976/1977 PDO shift (e.g., McGowan et al. 2003, Peabody et al. 2018), but our
560 analyses did not. Our estimated biology trend is consistent with the evidence of this regime shift,
561 however, only ichthyoplankton time series are available prior to the 1970s and there are gaps in
562 the ichthyoplankton data from the late 1960s through the 1970s. The trend estimate therefore has
563 higher uncertainty during this period than elsewhere in the time series.

564 The CCE biology time series included in this study showed strong coherence in
565 community signal in response to regional climate perturbations. Although they span across
566 multiple trophic levels, life-history strategies, and datasets, most of the biological time series
567 loaded in the same direction on the shared trend (Fig. 4b). In addition, our the CCE shared
568 biology trend and loadings captured an unusual aspect of the 2014–2016 warming events: the
569 abundance of several taxa, including young-of-year rockfish and anchovy, was high during the
570 marine heatwave (Santora et al. 2017, Thompson et al. 2019, Schroeder et al. 2019). By contrast,
571 their abundance was greatly reduced in most previous warm events, including two of the
572 strongest El Niño events on record (1982–1983, 1997–1998) and unusually low productivity

573 conditions (2005–2006, Peterson et al. 2006). High abundances of young-of-year rockfish and
574 groundfish, squid, and krill in the CCE are generally associated with more southward transport
575 and subarctic source waters, while abundances are typically far lower in years with more
576 subtropical waters, which are often associated with El Niño and anomalous warm events
577 (Ralston et al. 2015, Schroeder et al. 2019). The unexpectedly high abundance of these taxa in
578 2014–2016, despite surface-oriented marine heatwave, may be related to the prevalence of
579 subsurface waters that were more subarctic than subtropical in origin (Schroeder et al. 2019) and
580 to some concurrent strong upwelling events, particularly in spring 2015 (Peterson et al. 2015,
581 Ryan et al. 2017).

582 Our results were consistent with recent studies of several top predators in the CCE. The
583 DFA trends and loadings indicate a negative response of sea lion pup growth and weight to the
584 2014–2016 marine heatwave, which also aligns with past work showing that sea lion pup
585 condition covaries with abundance of forage such as larval anchovy and sardine, which provide
586 quality prey to sustain lactation in nursing mothers (McClatchie et al. 2016). Pup condition also
587 improved at the tail end of the marine heatwave when, despite the warm water, anchovy
588 abundance increased dramatically (Thompson et al. 2019). The trends and loadings suggest that
589 the reproductive success of some seabirds in the central CCE was not diminished by the
590 heatwave, although the heatwave had severe impacts seabird productivity in regions to the north
591 (Piatt et al. 2020).

592

593 *Regional comparison of the marine heatwave's effect on community state*

594 A compelling outcome of our analysis and a similar analysis applied to Alaskan species
595 by Litzow et al. (2020a) is that neither detected state changes in North Pacific communities

596 following the massive 2014–2016 marine heatwave, despite the extremely anomalous physical
597 conditions throughout most of the basin and a litany of concurrent biological, ecological, social
598 and economic effects (see Introduction). An important characteristic of both studies is the
599 temporal scale of community analysis (1972–2017 for the Gulf of Alaska (GOA) and 1951–2017
600 for the CCE). This long temporal scale provides an important context for comparing
601 contemporary change with the magnitude of historical community shifts. In addition, the
602 Bayesian DFA accounts for uncertainty in the shared trends in a way that prevents premature
603 detection of wholesale ecosystem shifts.

604 We note that Suryan et al. (2021), fitted a single-trend, non-Bayesian DFA model to a
605 larger set of GOA biological time series ($n = 187$) over a shorter time span (2010–2018) and
606 found evidence of a well-resolved shift that implied different community states during 2010–
607 2014 and 2015–2018. The different conclusions of Suryan et al. (2021) and Litzow et al. (2020a)
608 studies speak to an inherent tension in retrospective analyses of community change. Limited time
609 series availability means that analyses can be taxonomically and functionally broad (e.g., Suryan
610 et al. 2021), or temporally extensive (e.g., Litzow et al. 2020a), but not both. Each approach has
611 advantages, but direct comparison between the two is difficult. Given the impacts of the 2014–
612 2016 event, and that long-term warming combined with marine heatwaves will push the CCE
613 into novel climate states, we must consider ecological mechanisms that might explain why these
614 communities were apparently resilient to the marine heatwave, along with revisiting
615 methodological details that could further clarify our results.

616

617 *Environmental covariates*

618 DFA models of CCE biology that included a climate covariate performed better than
619 models without one. Nitrate flux into the surface mixed layer (BEUTI) was the best-performing
620 covariate for individual species in addition to the shared trend in the southern and central CCE
621 over the past three decades. Nitrate flux had a strong positive effect on the abundance of krill and
622 some larval fishes and on the reproductive success of seabirds, and a moderate positive effect on
623 some pelagic juvenile fishes, squid, and sea lion pup births. Stronger upwelling magnitude
624 (CUTI), which is correlated with nitrate flux, was the second-best predictor of community
625 variability and had a positive albeit weaker effect on the same suite of species (S4 Fig.). These
626 findings are consistent with mechanistic understanding as upwelling increases the supply of
627 nutrients to shallow waters and enhances the productivity at the lower trophic levels, including
628 juvenile rockfishes (Ralston et al. 2013), which affects foraging conditions for higher trophic
629 level species, such as seabirds (e.g., Wells et al. 2008).

630 BEUTI and CUTI had a strong, negative correlations with juvenile/adult Pacific sardine
631 and larval northern anchovy. The relative abundance of Pacific sardine in coastal waters off of
632 Central California has been shown to be lower during periods of strong upwelling (Santora et al.
633 2014, Ralston et al. 2015). This trend may reflect a change in the production of Pacific sardine or
634 a shift in their relative distribution. In addition, a negative relationship between upwelling and
635 sardine recruitment can generally be explained by the transfer of fish larvae to offshore areas
636 where they have low chance of survival during periods of strong equatorward flow and
637 upwelling (Bailey and Francis 1985, Nieto et al. 2014. Our understanding of the mechanisms
638 driving anchovy population dynamics is limited (Sydeman et al. 2020).

639 Climate drivers often act in concert to influence community variability, and here we are
640 evaluating the effects of climate variables one at a time. An important next step of this work will

641 be to examine whether including multiple climate covariates further improves the forecast skill
642 of the CCE biology and our community state indicator. However, the individual climate
643 variables are collinear and share information, which affects our ability to makes inference on the
644 covariates. Furthermore, our study is broad synthesis of community indicators and their response
645 to climate perturbations, and should not be interpreted as replacing more detailed investigations
646 into the drivers and mechanistic understanding of the indicators included here.

647

648 *Community state forecasting skill*

649 Our approach for creating simultaneous predictions of species responses and shared
650 ecosystem variability to ocean conditions shows promise for developing near-term forecasts of
651 community state. Our forecasts are based on outputs from the CCE ROMS, which have been
652 used to examine how oceanographic processes affect fish recruitment variability (Tolimieri et al.
653 2018, Haltuch et al. 2020) and productivity (Siegelman-Charbit et al. 2016), species habitat
654 suitability (Abrahms et al. 2018, Cimino et al. 2020), and species spatial distributions (Muhling
655 et al. 2019, 2020). The CCE ROMS also supports nowcasts of species distributions based on
656 observed ocean conditions (Hazen et al. 2017, 2018, Welch et al. 2019). Moreover, multiple
657 efforts are underway in the CCE and other coastal systems to use ROMS outputs to develop
658 short-term forecasts of ocean conditions for uptake by scientists, managers, and other end-users
659 (Siedlecki et al. 2016, Kaplan et al. 2016, Jacox et al. 2020, Malick et al. 2020). Here, we were
660 able to create forecasts of community state and several individual species parameters one year in
661 advance based on observations of a single climate variable (nitrate flux). Forecast lead times
662 could be extended further by using forecasts of ocean conditions rather than observed conditions,
663 and ocean temperatures in the CCE can be skillfully forecast months to a year in advance, with

664 particularly high skill in the late winter and spring (Jacox et al. 2019). Future extensions of our
665 work will evaluate whether different combinations of climate variables and time lags might
666 improve our forecasting skill.

667 Using DFA to forecast attributes of community structure in the CCE allows us to create
668 simultaneous forecasts of trends, or ‘ecosystem state’, and raw time series. Our approach could
669 also be applied individually to each dataset in our analysis to generate taxa-specific indicators
670 (e.g., seabird productivity, juvenile fish abundance), though these forecasts would be expected to
671 differ from those with the entire CCE dataset. Similarly, if ecosystem states were not a focus of
672 inference, alternative forecast models could be applied (e.g., ARIMA or non-parametric models,
673 Ward et al. 2014). Forecasts for individual time series from the DFA models used here can be
674 seen as a mixture of the AR forecast on the estimated trends (Fig. 6), and linear effects of
675 forecasted climate variables on each time series (Fig. 5). Species that have strong associations or
676 loadings on the trend and estimated climate effects that are large in magnitude (e.g., market
677 squid, Pacific sanddabs, shortbelly rockfish *Sebastodes jordani*) are expected to have the most
678 accurate predictions, while those species with weak loadings and weaker effects of climate
679 variables (e.g., California smoothtongue (*Leuroglossus stibius*) are expected to have poorer
680 forecast performance.

681 Nonstationary relationships are an important consideration for producing reliable
682 ecological forecasts. While the year-to-year variability in the estimated trend did appear to be
683 stationary in our community models (Fig. 4, 6), the autocorrelation appeared to be nonstationary
684 with the lag-1 autocorrelation between 2000–present being significantly higher (0.82) than over
685 the years 1981–2000 (0.23). In addition to nonstationary variance parameters, future analyses
686 may also consider nonstationary relationships in the covariate relationships, or potential

687 interactions between covariates. A growing number of retrospective analyses have revealed
688 nonstationary relationships among climate and individual species or community-level variables
689 (Puerta et al. 2019, Litzow et al. 2018, 2019, 2020 a,b,c). In the northeast Pacific Ocean, these
690 studies have been mostly focused on Alaskan ecosystems with long time series describing
691 climate and biological processes. The best-documented instance of nonstationary relationships
692 among climate and biology time series in the North Pacific centers on a climate shift in the late
693 1980s (Litzow et al. 2020b). Decades of observational data on either side of that event allow for
694 statistically robust tests for nonstationarity that are not yet available for post-2014–2016
695 conditions. However, early indications from Alaska suggest the possibility that long-standing
696 relationships between leading climate modes and individual climate and biology time series may
697 have changed following 2014 (Litzow et al. 2020c).

698

699 *Management application*

700 Our approach for developing a community state indicator to track and predict the
701 response of marine ecosystems to climate perturbations has the potential to support ecosystem-
702 based and climate-ready management in multiple ways. Garnering knowledge of community
703 state and the potential for large shifts in ecosystem structure in response to intense and novel
704 climate perturbations can help inform better, more rapid management decisions for mitigating
705 ecological and socioeconomic impacts. Our intention is to continually update our analyses when
706 new data become available to provide the most up-to-date information on the CCE community
707 state for scientists, managers, and stakeholders.

708 The combination of long-term monitoring surveys and data with the modeling framework
709 we advance here can also help scientists identify or refine key variables of ecosystem change that

710 are summarized for ecosystem assessments in support of decision-making (Harvey et al. 2020).
711 For example, it might be prudent to emphasize ecological time series that load strongest on
712 ecosystem state trends and demonstrate strong, predictable relationships with climate variables
713 (or other covariates of interest) over time series with weaker loadings or lower forecast skill.
714 Furthermore, our approach can provide valuable ecosystem information for scientific,
715 management and coastal communities during times when researchers cannot sample the biology
716 in marine ecosystems. This added value became acutely apparent in 2020 when myriad ocean
717 surveys were cancelled or limited in spatiotemporal scope due to safety restrictions associated
718 with the COVID-19 pandemic.

719 Finally, our approach provides a quantitative way to help managers discern short-term
720 periods of unusual community dynamics and/or high variability—such as the 2014–2016 marine
721 heatwave—from state shifts that represent more enduring transitions into new regimes of
722 ecosystem structure or productivity. Given that global climate change is expected to amplify
723 ocean change, approaches like the one applied here will become increasingly valuable for
724 identifying novel community states that require new marine resource management and
725 conservation considerations.

726

727 **Acknowledgements**

728 We thank those who have spent countless hours planning field surveys and collecting data for the
729 invaluable time series used in our study. We thank the U.S. Fish and Wildlife Service for
730 granting permission and providing resources to conduct research on the Farallon Islands National
731 Wildlife Refuge. Funders for Point Blue’s Farallon Research Program include the Bently
732 Foundation, Elinor Patterson Baker Trust, Bernice Barbour Foundation, Frank A. Campini

733 Foundation, Grand Foundation, Kimball Foundation, Marisla Foundation, Giles W. and Elise G.
734 Mead Foundation, Moore Family Foundation, RHE Charitable Foundation, Volgenau
735 Foundation, and numerous individual donors. We thank B. Feist for creating the map of the
736 sampling areas (Figure 1). We thank N. Tolimieri and T.L. Rogers for their helpful comments on
737 an earlier version of this manuscript. We thank the Journal's academic editor, Wei Yu, and two
738 reviewers, Shuyang Ma and Chongliang Zhang, for their constructive comments that improved
739 our manuscript. Funding for this project came from NOAA's Fisheries and the Environment
740 (FATE) program (project 16-01) awarded to M.E.H, E.J.W., M.A.L. and C.J.H. and NOAA's
741 California Current Integrated Ecosystem Assessment program (C.J.H.). The funders had no role
742 in study design, data collection and analysis, decision to publish, or preparation of the
743 manuscript.

744

745

746 **References**

- 747 Abrahms B, Welch H, Brodie S, Jacox MG, Becker EA, Bograd SJ et al. Dynamic ensemble
748 models to predict distributions and anthropogenic risk exposure for highly mobile species.
749 Divers. Distrib. 2019; 25: 1182-1193. doi: 10.1111/ddi.12940.
- 750 Anderson SC, Ward EJ. Black swans in space: modelling spatiotemporal processes with
751 extremes. Ecology 2019; 100: e02403. doi:10.1002/ecy.2403.
- 752 Anderson SC, Branch TA, Cooper AB, Dulvy NK. Black-swan events in animal populations.
753 Proc. Natl. Acad. Sci. U.S.A. 2017; 114: 3252–3257.
- 754 Anderson PJ, Piatt JF. Community reorganization in the Gulf of Alaska following ocean climate
755 regime shift. Mar. Ecol. Prog. Ser. 1999; 189: 117-123.

- 756 Amante C, Eakins BW.ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources
757 and Analysis. NESDIS NGDC-24, National Geophysical Data Center, United States
758 Department of Commerce, Boulder, CO, March 2009. 19 p. DOI: 10.7289/V5C8276M
- 759 Bailey KM, Francis RC. Recruitment of Pacific Whiting, *Merluccius productus*, and the ocean
760 environment. Mar. Fish. Rev. 1985; 47: 8-11.
- 761 Baumgartner TR, Soutar A, Ferreira-Bartrina V. Reconstructions of the history of Pacific sardine
762 and northern anchovy populations over the past two millennia from sediments of the Santa
763 Barbara Basin, California. CalCOFI Rep. 1992; 33: 24-40.
- 764 Beaugrand G, Edwards M, Brander K, Luczak C, Ibanez F. Causes and projections of abrupt
765 climate-driven ecosystem shifts in the North Atlantic. Ecol. Lett. 2008; 11: 1157–1168.
- 766 Beaugrand G, Conversi A, Chiba S, Edwards M, Fonda-Umani S, Greene C et al. Synchronous
767 marine pelagic regime shifts in the Northern Hemisphere. Phil. Trans. R. Soc. 2016; 370:
768 20130272. doi: 10.1098/rstb.2013.0272.
- 769 Benson AJ, Trites AW. Ecological effects of regime shifts in the Bering Sea and eastern North
770 Pacific Ocean. Fish Fish. 2002; 3: 95–113.
- 771 Bjorkstedt EP, Goericke R, McClatchie S, Weber E, Watson W, Lo N et al. State of the
772 California current 2009-2010: Regional variation persists through transition from the la
773 Niña to el Niño (and back?). CalCOFI Rep. 2010; 51.
- 774 Bograd S, Schroeder ID, Jacox MG. A water mass history of the Southern California current
775 system. Geophys. Res. Lett. 2019; 46: 6690-6698.
- 776 Bond NA, Cronin MF, Freeland H, Mantua N. Causes and impacts of the 2014 warm anomaly in
777 the NE Pacific. Geophys. Res. Lett. 2015; 42, 3414–3420. doi: 10.1002/2015GL063306.
- 778 Bürkner P-C, Gabry J, Vehtari A. Approximate leave-future-out cross-validation for Bayesian

- 779 time series models. *J. Stat Comput. Simul.* 2020; 90: 2499-2523.
- 780 Brodeur RD, Auth TD, Phillips AJ. Major shifts in pelagic micronekton and zooplankton
781 community structure in an upwelling ecosystem related to an unprecedented marine
782 heatwave. *Front. Mar. Sci.* 2019. doi: 10.3389/fmars.2019.00212.
- 783 Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M et al. Stan: A
784 Probabilistic Programming Language. *J. Stat. Softw.* 2017. doi: 0.18637/jss.v076.i01.
- 785 Cavole LM, Demko AM, Diner RE, Giddings A, Koester I, Pagniello CM et al. Biological
786 impacts of the 2013–2015 warm-water anomaly in the Northeast Pacific: winners, losers,
787 and the future. *Oceanography* 2016; 29, 273–285.
- 788 Chavez FP, Ryan J, Lluch SE, Niquen MC. From Anchovies to Sardines and Back: Multidecadal
789 change in the Pacific Ocean. *Science* 2003; 299: 217-221.
- 790 Checkley DM Jr, Barth JA. Patterns and processes in the California Current System. *Prog.*
791 *Ocean.* 2009; 83, 49-64.
- 792 Cimino MA, Santora JA, Schroeder I, Sydeman W, Jacox MG, Hazen EL, Bograd SJ. Essential
793 krill species habitat resolved by seasonal upwelling and ocean circulation models within the
794 large marine ecosystem of the California Current System. *Ecography* 2003; 43, 1-15. doi:
795 10.1111/ecog.05204.
- 796 Cohn TA, Lins, HF. Nature's style: Naturally trendy. *Geophys. Res. Lett.* 2005. doi:
797 10.1029/2005GL024476
- 798 Deyle ER, Fogarty M, Hsieh C-h, Kaufman L, MacCall AD, Munch SB et al. Predicting climate
799 effects on Pacific sardine. *Proc Natl Acad Sci USA* 2013; 110: 6430-6435.
- 800 Field DB, Baumgartner TR, Ferreira V, Gutierrez D, Lozano-Montes H, Salvatteci R, Soutar A.
801 Variability from scales in marine sediments and other historical records. In: Checkley DM,

- 802 Alheit J, Oozeki Y, editors. Climate change and small pelagic fish. Cambridge: Cambridge
803 University Press; 2009. pp. 45-63.
- 804 Field JC, Miller RA, Santora JA, Tolimieri N, Haltuch MA, Brodeur RD et al. Spatiotemporal
805 patterns of variability in the abundance and distribution of winter-spawned pelagic juvenile
806 rockfish in the California Current. PloS one 2021; 16: e0251638.
- 807 Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis. 3rd
808 ed. CRC Press; 2013.
- 809 Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. Statist.
810 Sci. 1992; 7: 457-472.
- 811 Gottscho AD. Zoogeography of the San Andreas Fault system: Great Pacific Fracture Zones
812 correspond with spatially concordant phylogeographic boundaries in western North
813 America. Biological Reviews 2016; 91: 235-254.
- 814 Haltuch MA, Tolimieri N, Lee Q, Jacox MG. Oceanographic drivers of petrale sole recruitment
815 in the California Current Ecosystem, Fish. Ocean. 2020; 29:122-136. doi:
816 10.1111/fog.12459.
- 817 Hare SR, Mantua NJ. Empirical evidence for North Pacific regime shifts in 1977 and 1989. Prog.
818 Oceanogr. 2000; 47: 103–145. doi: 10.1016/s0079-6611(00)00033-1.
- 819 Harvey CJ, Fisher J, Samhouri, JF, Williams GD, Francis TB, Jacobson KC et al. The
820 importance of long-term ecological time series for integrated ecosystem assessment and
821 ecosystem-based management. Prog. Oceanogr. 2020. doi: 10.1016/j.pocean.2020.102418.
- 822 Hazen EL, Palacios DM, Forney KA, Howell EA, Becker E, Hoover AL et al. WhaleWatch: a
823 dynamic management tool for predicting blue whale density in the California Current. J.
824 Appl. Ecol. 2017. doi: 10.1111/1365-2664.12820.

- 825 Hazen EL, Scales KL, Maxwell SM, Briscoe D, Welch, H, Bograd et al. A dynamic ocean
826 management tool to reduce bycatch and support sustainable fisheries. *Sci. Adv.* 2018; 4:
827 eaar3001.
- 828 Hobday AJ, Oliver ECJ, Sen Gupta A, BenthuySEN JA, Burrows MT, Donat MG et al.
829 Categorizing and naming marine heatwaves. *Oceanography* 2018; 31: 162 - 173.
- 830 Hobday AJ, Spillman CM, Paige Eveson J, Hartog JR. Seasonal forecasting for decision support
831 in marine fisheries and aquaculture. *Fish. Oceanogr.* 2016; 25: 45–56.
- 832 Hoffman MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in
833 Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* 2014; 15: 1593–1623.
- 834 Holmes EE. Beyond theory to applications and evaluation: Diffusions approximations for
835 population viability analysis. *Ecol. Appl.* 2004; 14: 1272-1293.
- 836 Holmes EE, Ward EJ, Scheuerell, MD. Analysis of multivariate time series using the MARSS
837 package; 2018.
- 838 Jacox MG, Alexander MA, Siedlecki S, Chen K, Kwon Y-O, Brodie S et al. Seasonal-to-
839 interannual prediction of North American coastal marine ecosystems: Forecast methods,
840 mechanisms of predictability, and priority developments. *Prog. Oceanogr.* 2020. doi:
841 10.1016/j.pocean.2020.102307.
- 842 Jacox MG, Alexander MA, Stock CA, Hervieux G. On the skill of seasonal sea surface
843 temperature forecasts in the California Current System and its connection to ENSO
844 variability. *Clim. Dyn.* 2019; 53: 7519-7533.
- 845 Jacox MG, Alexander MA, Mantua NJ, Scott JD, Hervieux G. Webb RS et al. Forcing of
846 multiyear extreme ocean temperatures that impacted California Current living marine
847 resources in 2016. *Bull. Am. Meteorol. Soc.* 2018a; 99: S27–S33. doi: 10.1175/BAMS-D-

- 848 17-0119.1
- 849 Jacox MG, Edwards CA, Hazen EL, Bograd, SJ. Coastal upwelling revisited: Ekman, Bakun,
850 and improved upwelling indices for the U.S. west coast. *J. Geophys. Res.* 2018b.
851 doi:10.1029/2018JC014187.
- 852 Jacox MG, Fletcher J, Moore AM, Edwards CA. ENSO and the California Current coastal
853 upwelling response. *J. Geophys. Res.* 2015. doi:10.1002/2014JC010650
- 854 Jones T, Parish JK, Peterson WT, Bjorkstedt EP, Bond NA, Balance LT et al. Massive mortality
855 of a planktivorous seabird in response to a marine heatwave. *Geophys. Res. Lett.* 2018; 45:
856 3193-3202.
- 857 Kaplan IC, Williams GD, Bond NA, Hernmann AJ, Siedlecki S. et al. Cloudy with a chance of
858 sardines: forecasting sardine distributions using regional climate models. *Fish. Oceanogr.*
859 2016; 25: 15-2.
- 860 Koslow J, Goericke R, Watson W. Fish assemblages in the Southern California Current:
861 relationships with climate, 1951–2008. *Fish. Oceanogr.* 2013; 22: 207–219.
- 862 Koslow JA, Hobday AJ, Boehlert GW et al. Climate variability and marine survival of coho
863 salmon (*Oncorhynchus kisutch*) in the Oregon production area. *Fish. Oceanogr.* 2002; 11: 65-
864 77.
- 865 Laufkötter C, Zscheischler J, Frölicher TL. High-impact marine heatwaves attributable to human-
866 induced global warming. *Science* 2020; 369: 1621-1625.
- 867 Litzow MA, Ciannelli L. 2007. Oscillating trophic control induces community reorganization in
868 a marine ecosystem. *Ecol. Lett.* 2007; 10: 1124–1134. doi: 10.1111/j.1461-
869 0248.2007.01111.x.
- 870 Litzow MA, Ciannelli L, Puerta P, Wettstein JJ, Rykaczewski RR, Opiekun M. Non-stationary

871 climate–salmon relationships in the Gulf of Alaska. Proc. R. Soc. B Biol. Sci. 2018; 285:
872 20181855. doi: 10.1098/rspb.2018.1855.

873 Litzow MA, Ciannelli L, Puerta P, Wettstein JJ, Rykaczewski RR, Opiekun M. Nonstationary
874 environmental and community relationships in the North Pacific Ocean. Ecology 2019; 100:
875 ecy.2760. doi: 10.1002/ecy.2760.

876 Litzow MA, Hunsicker ME, Ward EJ, Anderson SC, Gao J, Zador S et al. Evaluating ecosystem
877 change as Gulf of Alaska temperature exceeds the limits of preindustrial variability. Prog.
878 Ocean. 2020a; 117: 7665–7671.

879 Litzow MA, Hunsicker ME, Bond NA, Burke BJ, Cunningham CJ, Gosselin JL, Norton EL,
880 Ward EJ, Zador SG. The changing physical and ecological meanings of North Pacific Ocean
881 climate indices. Proc. Natl. Acad. Sci. U.S.A. 2020b; 117: 7665–7671. doi:
882 10.1073/pnas.1921266117.

883 Litzow MA, Malick MJ, Bond NA, Cunningham CJ, Gosselin JL, Ward EJ. Quantifying a novel
884 climate through changes in PDO-climate and PDO-salmon relationships. Geophys. Res. Lett.
885 2020c. doi: 10.1029/2020GL087972.

886 MacCall AD. Patterns of low-frequency variability in fish populations of the California Current.
887 CalCOFI Reports 1996; 37: 100–110.

888 Malick M.J, Siedlecki SA, Norton EL, Kaplan IC, Haltuch MA, Hunsicker ME et al.
889 Environmentally driven seasonal forecasts of Pacific hake distribution. Front. Mar. Sci.
890 2020; 7: 578490. doi: 10.3389/fmars.2020.578490.

891 Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC. A Pacific interdecadal climate
892 oscillation with impacts on salmon production. Bull. Am. Meteorol. Soc. 1997; 78, 1069–
893 1079.

- 894 McCabe RM, Hickey BM, Kudela RM, Lefebvre KA, Adams NG, Bill BD et al. An
895 unprecedeted coastwide toxic algal bloom linked to anomalous ocean conditions. Geophys.
896 Res. Lett. 2016; 43: 10366–10376. doi:10.1002/2016GL070023.
- 897 McClatchie S, Goericke R, Koslow JA, Schwing FB, Bograd SJ, Charter RW et al. The State of
898 the California Current, 2007–2008: La Niña conditions and their effects on the ecosystem.
899 Cal-COFI Rep. 2008; 49: 39–76.
- 900 McClatchie S, Goericke R, Schwing FB, Bograd SJ, Peterson WT, Emmett R et al. The state of
901 the California Current, 2008–2009: Cold conditions drive regional difference. CalCOFI
902 Rep. 2009; 50: 43–68.
- 903 McClatchie S, Duffy-Anderson J, Field JC, Goericke R, Griffith D, Hanisko DS, Hare JA,
904 Lyczkowski-Shultz J, Peterson WT, Watson W, Weber ED. Long time series in US fisheries
905 oceanography. Oceanography 2014; 27: 48-67.
- 906 McClatchie S, Field J, Thompson AR, Gerrodette T, Lowry M, Fiedler PC et al. Food limitation
907 of sea lion pups and the decline of forage off central and southern California. R. Soc.
908 opensci. 2016; 3: 150628. doi: 10.1098/rsos.150628.
- 909 McGowan JA, Bograd SJ, Lynn RJ, Miller AJ. The biological response to the 1977 regime shift
910 in the California Current. Deep Sea Res. II 2003; 50: 2567-2582.
- 911 Milly PCD, Betancourt J, Falkenmark M, Hirsch, RM, Kundzewicz ZW, Lettenmaier DP et al.
912 On critiques of “Stationarity is dead: Whither water management?” Water Res. Res. 2015;
913 51: 7785–7789.
- 914 Möllmann C, Diekmann R. Marine Ecosystem Regime Shifts Induced by Climate and
915 Overfishing: A Review for the Northern Hemisphere. Adv. Ecol. Res. 2012; 47: 303.347.
- 916 Morgan CA, Beckman BR, Weitkamp LA, Fresh KL. Recent ecosystem disturbance in the

- 917 Northern California Current. *Fisheries* 2019; 44: 465-474. doi: 10.1002/fsh.10273.
- 918 Muhling B, Brodie S, Jacox MG, Snodgrass O, Dewar H, Tommasi D, Edwards C, Xu Y, Snyder
919 S, Childers J. Dynamic habitat use of albacore and their primary prey species in the
920 California Current System. *CalCOFI Reports* 2019; 60: 79-93.
- 921 Muhling B, Brodie S, Smith JA, Tommasi D, Gaitan CF, Hazen EL et al. Predictability of
922 species distributions deteriorates under novel environmental conditions in the California
923 Current System. *Front. Mar. Sci.* 2020; doi:10.3389/fmars.2020.00589.
- 924 Neveu E, Moore AM, Edwards CA, Fiechter J, Drake P, Jacox MG, Nuss E. A historical analysis
925 of the California Current using ROMS 4D-Var. Part I: System configuration and
926 diagnostics, *Ocean Model.* 2016; 99: 133-151. doi:10.1016/j.ocemod.2015.11.012.
- 927 Nielsen JM, Rogers LA, Brodeur RD, Thompson AR, Auth TD, Dreary AL et al. Responses of
928 ichthyoplankton assemblages to the recent marine heatwave and previous climate
929 fluctuations in several Northeast Pacific marine ecosystems. *Glob. Chang. Biol.* 2020; 27:
930 506-520.
- 931 Nieto K, McClatchie S, Weber ED, Lennert-Cody CE. Effect of mesoscale eddies and streamers
932 on sardine spawning habitat and recruitment success off Southern and central California, *J.*
933 *Geophys. Res. Oceans* 2014; 119: 6330–6339, doi:10.1002/2014JC010251.
- 934 Peabody CE, Thompson AR, Sax DF, Morse RE, Perretti CT. Decadal regime shifts in southern
935 California's ichthyoplankton assemblage. *Mar. Ecol. Prog. Ser.* 2018; 607: 71-83.
- 936 Peterson WT, Emmett R, Goericke R, Venrick E, Mantyla A, Bograd S et al. The State of the
937 California Current, 2005-2006: warm in the north, cool in the south. *CalCOFI Reports*
938 2006; 47: 30–74.
- 939 Peterson WT, Fisher JL, Strub PT, Du X, Risien C, Peterson J et al. The pelagic ecosystem in

940 the Northern California Current off Oregon during the 2014-2016 warm anomalies within
941 the context of the past 20 years. *J. Geophys. Res.* 2017; 122: 7267-7290.

942 Piatt JF, Parrish JK, Renner HM, Schoen SK, Jones TT, Arimitsu ML et al. Extreme mortality
943 and reproductive failure of common murres resulting from the northeast Pacific marine
944 heatwave of 2014-2016. *PLoS ONE* 2020; 15: e0226087. doi:
945 10.1371/journal.pone.0226087

946 Planque B, Arneberg P. Principal component analyses for integrated ecosystem assessments may
947 primarily reflect methodological artefacts. *ICES J. Mar. Sci.* 2018; 75:
948 1021–1028.

949 Puerta P, Ciannelli L, Rykaczewski R, Opiekun M, Litzow MA. Do Gulf of Alaska fish and
950 crustacean populations show synchronous non-stationary responses to climate? *Prog.*
951 *Oceanogr.* 2019; 175: 161–170. doi: 10.1016/j.pocean.2019.04.002.

952 R Core Team. R: A language and environment for statistical computing. R Foundation for
953 Statistical Computing, Vienna, Austria, 2021.

954 Ralston S, Sakuma KM, Field JC. Interannual variation in pelagic juvenile rockfish (*Sebastes*
955 spp.) abundance – going with the flow. *Fish. Ocean.* 2013; 22:288–308. doi:
956 10.1111/fog.12022.

957 Ralston S, Field JC, Sakuma KM. Long-term variation in a central California pelagic forage
958 assemblage. *J. Mar. Sys.* 2015; 146: 26-37. doi: 10.1016/j.jmarsys.2014.06.013.

959 Ryan JP, Kudela RM, Birch JM, Blum M, Bowers HA, Chavez FP et al. Causality of an extreme
960 harmful algal bloom in Monterey Bay, California, during the 2014–2016 northeast Pacific
961 warm anomaly. *Geophys. Res. Lett.* 2017; 44: 5571-5579. doi: 10.1002/2017GL072637.

- 962 Sakuma KM, Field JC, Mantua NJ, Ralston S, Marinovic BB, Carrion CN. Anomalous
963 epipelagic micronekton assemblage patterns in the neritic waters of the California Current
964 in spring 2015 during a period of extreme ocean conditions. CalCOFI Reports 2016; 57:
965 163-183.
- 966 Sanford E, Sones JL, García-Reyes M, Goddard JH, Largier JL. Widespread shifts in the coastal
967 biota of northern California during the 2014–2016 marine heatwaves. Sci. Rep. 2019; 9: 1-
968 14.
- 969 Santora JA, Mantua NJ, Schroeder ID, Field JC, Hazen E, Bograd SJ et al. Habitat compression
970 and ecosystem shifts as potential links between marine heatwave and record whale
971 entanglements. Nat. Commun. 2020; 11: 1-12.
- 972 Santora JA, Hazen EL, Schroeder ID, Bograd SJ, Sakuma KM, Field JC. Impacts of ocean
973 climate variability on biodiversity of pelagic forage species in an upwelling ecosystem.
974 Mar. Ecol. Prog. Ser. 2017; 580: 205-220.
- 975 Santora JA, Schroeder ID, Field JC, Wells BK, Sydeman WJ. Spatiotemporal dynamics of ocean
976 conditions and forage taxa reveals regional structuring of predator-prey relationships. Ecol.
977 Appl. 2014; 24:1730-1747. doi:10.1890/13-1605.1.
- 978 Schroeder ID, Santora JA, Bograd SJ, Hazen EL, Sakuma, KM, Moore AM et al. Source water
979 variability as a driver of rockfish recruitment in the California Current Ecosystem:
980 implications for climate change and fisheries management. Can. J. Fish. Aquat. Sci. 2019;
981 76: 950-960. doi: 10.1139/cjfas-2017-0480.
- 982 Schwartzlose RA, Alheit J, Bakun A, Baumgartner TR, Cloete R, Crawford RJM et al.
983 Worldwide large-scale fluctuations of sardine and anchovy populations. S. Afr. J. Mar. Sys.
984 1999; 21: 289-347.

- 985 Sen Gupta A, Thomsen M, Benthuysen JA, Hobday AL, Oliver E, Alexander LV et al. Drivers
986 and impacts of the most extreme marine heatwaves events. *Sci. Rep.* 2020; 10: 19359. doi:
987 10.1038/s41598-020-75445-3.
- 988 Seo H, Brink KH, Dorman E, Koracin D, Edwards CA. What determines the spatial pattern in
989 summer upwelling trends on the US West Coast? *J. Geophys. Res. Oceans* 2012. doi:
990 10.1029/2012JC008016.
- 991 Siedlecki SA, Kaplan IC, Hermann AJ, Nguyen TT, Bond NA, Newton JA et al. Experiments
992 with seasonal forecasts of ocean conditions for the northern region of the California
993 Current upwelling system. *Sci. Rep.* 2016; 6: 27203. DOI: 10.1038/srep27203.
- 994 Siegelman-Charbit L, Koslow JA, Jacox MG, Hazen EL, Bograd SJ, Miller EF, McGowan JA.
995 Physical forcing on fish abundance in the southern California Current System. *Fish. Ocean.*
996 2018; 27: 475–488. doi:10.1111/fog.12267.
- 997 Stan Development Team. RStan: The R interface to Stan. 2018.
- 998 Suryan RM, Arimitsu ML, Coletti HA, Hopcroft RR, Lindeberg MR, Barbeaux SJ et al.
999 Ecosystem response persists after a prolonged marine heatwave. *Sci. Rep.* 2021; 11: 6235.
- 1000 Sydeman WJ, Dedman S, Garcia-Reyes M, Thompson SA, Thayer JA, Bakun A, MacCall AD.
1001 Sixty-five years of northern anchovy population studies in the southern California Current:
1002 a review and suggestion for sensible management. *ICES J. Mar. Sci.* 2020; 77: 486–499.
- 1003 Thompson AR, Hyde JR, Watson W, Chen DC, Guo LW. Rockfish assemblage structure and
1004 spawning locations in southern California identified through larval sampling. *Mar. Ecol.*
1005 *Prog. Ser.* 2016; 547: 177-192.

- 1006 Thompson AR, Schroeder ID, Bograd SJ, Hazen EL, Jacox MG, Leising AL et al. State of the
1007 California Current 2018-19: a novel anchovy regime and a new marine heatwave?
1008 CalCOFI Reports 2019; 60: 1-65.
- 1009 Tolimieri N, Haltuch M, Lee Q, Jacox MG, Bograd SJ. Oceanographic drivers of sablefish
1010 recruitment in the California Current. Fish. Ocean, 2018; 27: 458-474,
1011 doi:10.1111/fog.12266.
- 1012 Tommasi D, Stock CA, Hobday AJ, Methot R, Kaplan IC, Eveson JP et al. Managing living
1013 marine resources in a dynamic environment: the role of seasonal to decadal climate
1014 forecasts. Prog. Oceanogr. 2017; 152: 15–49.
- 1015 Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-
1016 validation and WAIC. Statistics and Computing 2017; 27: 1413-1432. doi:
1017 10.1007/s11222-016-9696-4.
- 1018 Walker Jr HJ, Hastings PA, Hyde JR, Lea RN, Snodgrass OE, Bellquist LF. Unusual occurrences
1019 of fishes in the Southern California Current System during the warm water period of 2014–
1020 2018. Estuar. Coast. Shelf Sci. 2020; 236: 106634.
- 1021 Walsh JE, Thoman RL, Bhatt US, Bieniek PA, Brettschneider B, Brubaker M et al. The high
1022 latitude heat wave of 2016 and its impacts on Alaska. Bull. Am. Meteorol. Soc. 2018; 99:
1023 S39–S43. doi: 10.1175/BAMS-D-17-0105.
- 1024 Ward EJ, Holmes EE, Thorson JT, Collen B. Complexity is costly: a meta-analysis of parametric
1025 and non-parametric methods for short-term population forecasting. Oikos 2014; 123: 652–
1026 661.
- 1027 Ward EJ, Anderson SC, Damiano LA, Hunsicker ME, Litzow MA. Modeling regimes with
1028 extremes: the bayesdfa package for identifying and forecasting common trends and

1029 anomalies in multivariate time-series data. *R J* 2019; 11: 46–55.

1030 Ward EJ, Anderson SC, Damiano LA, Malick MJ. *bayesdfa: Bayesian Dynamic Factor Analysis*
1031 (DFA) with 'Stan'. R package version 1.1.0 2020. [https://CRAN.R-
1032 project.org/package=bayesdfa](https://CRAN.R-project.org/package=bayesdfa)

1033 Welch H, Hazen EL, Briscoe DK, Bograd SJ, Jacox MG, Eguchi T et al. Environmental
1034 indicators to reduce loggerhead turtle bycatch offshore of Southern California. *Ecol. Ind.*
1035 2019; 98: 657-664. doi:10.1016/j.ecolind.2018.11.001.

1036 Wells BK, Field JC, Thayer JA, Grimes CB, Bograd SJ, Sydeman WJ et al. Untangling the
1037 relationships among climate, prey and top predators in an ocean ecosystem. *Mar. Ecol.
1038 Prog. Ser.* 2008; 364: 15-29.

1039 Wernberg T, Bennett S, Babcock RC, De Bettignies T, Cure K, Depczynski M et al. Climate-
1040 driven regime shift of a temperate marine ecosystem. *Science* 2016; 353: 169-172.

1041 Zuur AF, Tuck ID, Bailey N. Dynamic factor analysis to estimate common trends in fisheries
1042 time series. *Can. J. Fish. Aquat. Sci.* 2003; 60: 542–552.

1043

1044

1045 **Supporting Information**

1046 **S1 Appendix:** Standardization of time series from spatially resolved datasets.

1047 **S1 Table:** Climate and biology time series included in the analyses

1048 **S2 Table:** Summary information for the Bayesian DFA biology-covariate and biology only
1049 models (years 1981–2017).

1050 **S3 Table:** Observations, predictions, and prediction errors for single species parameters in 2018.

1051 **S1 Figure:** Climate and biology time series used in the study analyses.

1052 **S2 Figure:** AR(1) coefficient on the southern/central California latent climate trend and support
1053 for a heavy-tailed deviations of the latent trend.

1054 **S3 Figure:** The Student-t deviations degrees of freedom parameter (nu) in the southern/central
1055 California biology trend.

1056 **S4 Figure:** A summary of the effect of the Cumulative Upwelling Transport Index (CUTI) on
1057 the individual single species parameter included in the DFA analyses

1058 **S5 Figure:** A summary of the effect of the Isothermal Layer Depth (ILD) on the individual
1059 single species parameter included in the DFA analyses.

1060 **S6 Figure:** A summary of the effect of the sea surface temperature on the individual single
1061 species parameter included in the DFA analyses

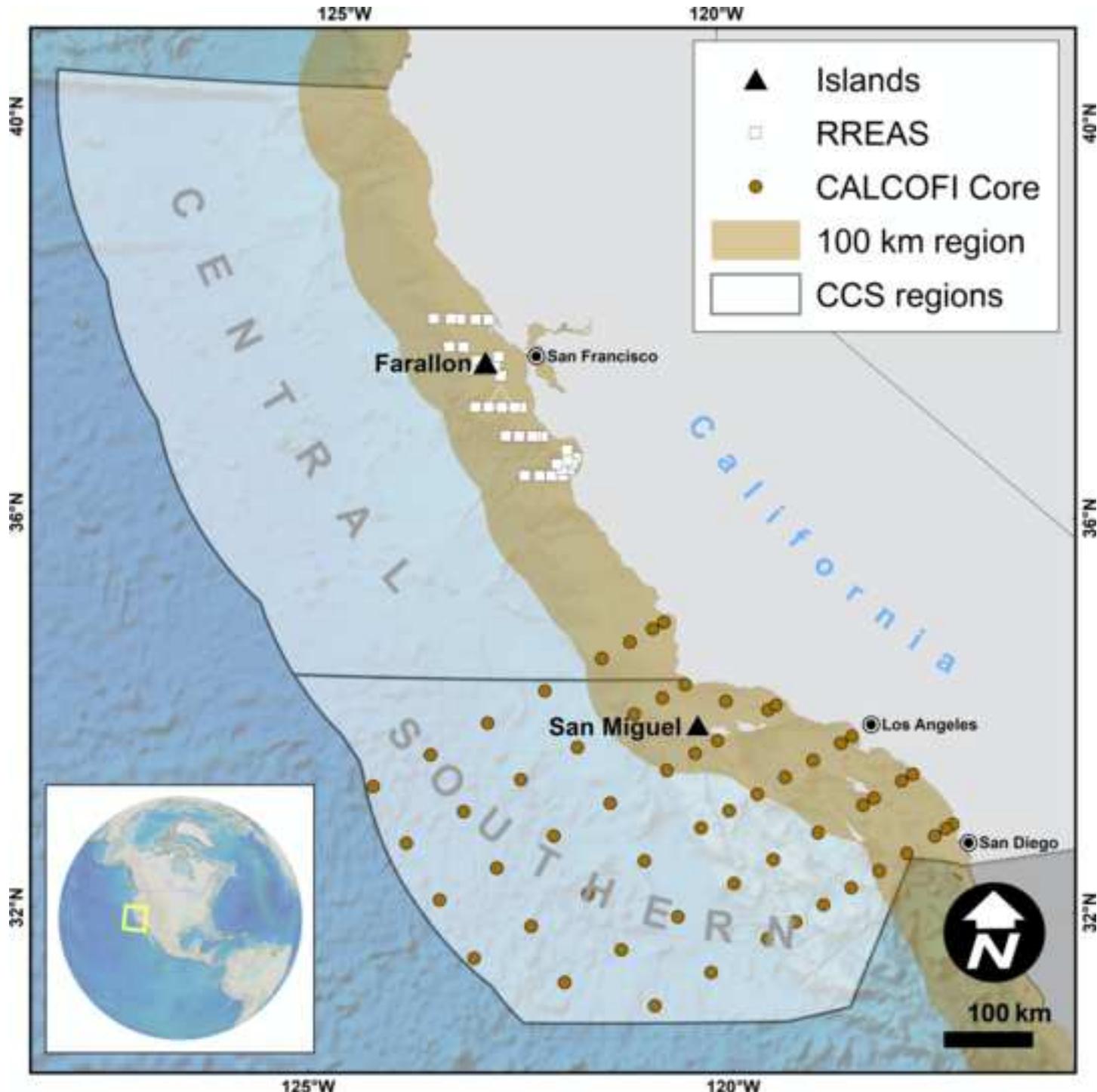
1062 **S7 Figure:** Community variability in the southern California Current ecosystem (1981–2018).

1063 **S8 Figure:** Fitted values for biology-covariate model including BEUTI (nitrate flux) as a
1064 covariate (1981–2017).

1065 **S9 Figure:** Log coefficient of variation (CV) of 2018 predictions of individual species
1066 parameters plotted against the mean and log CV of loadings related to each species, and the
1067 mean and log CV of coefficients relating each species to BEUTI (nitrate flux).

Figure 1 Study map

Click here to access/download;Figure;Figure_1_CCE study area -revised-open source.tif



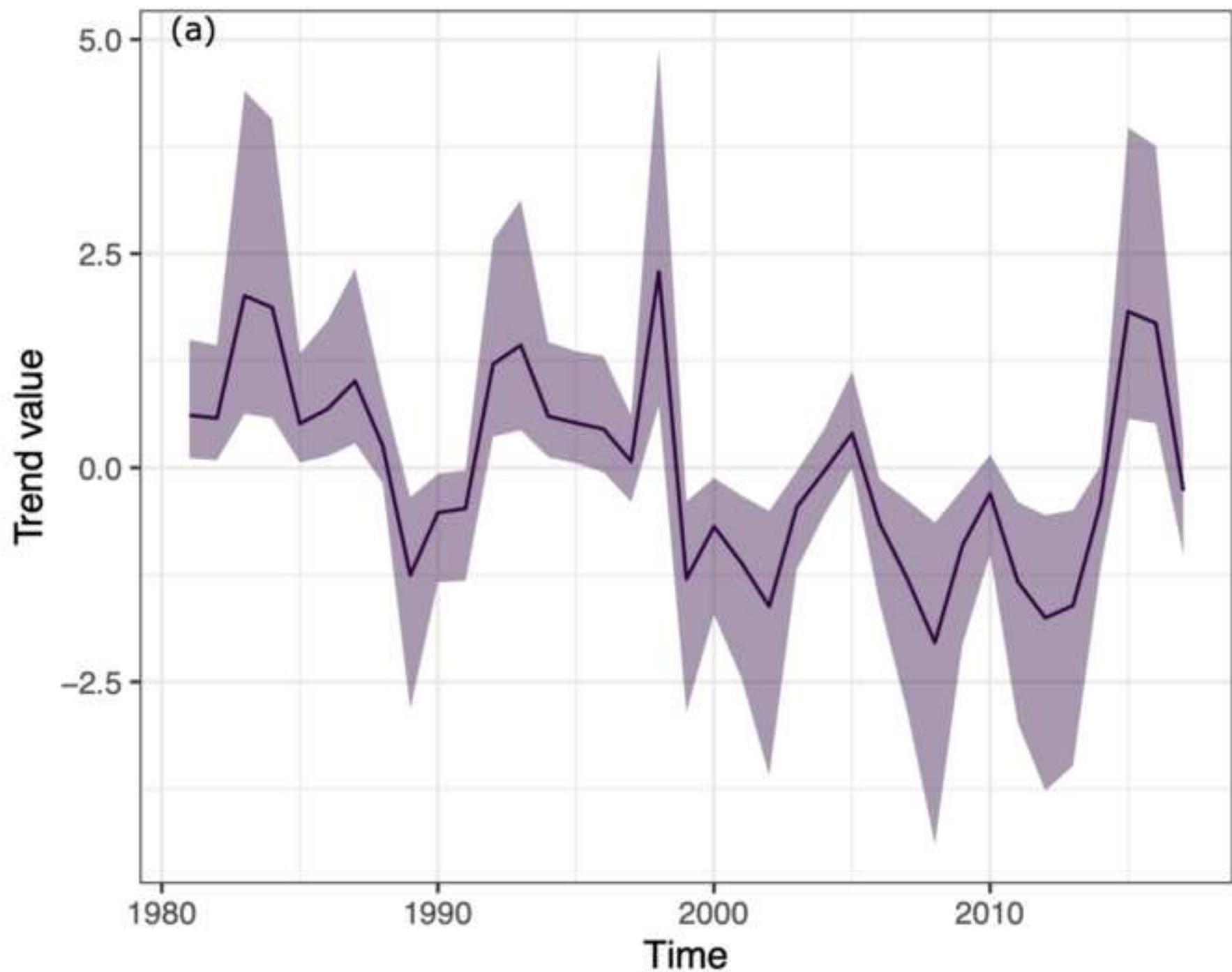
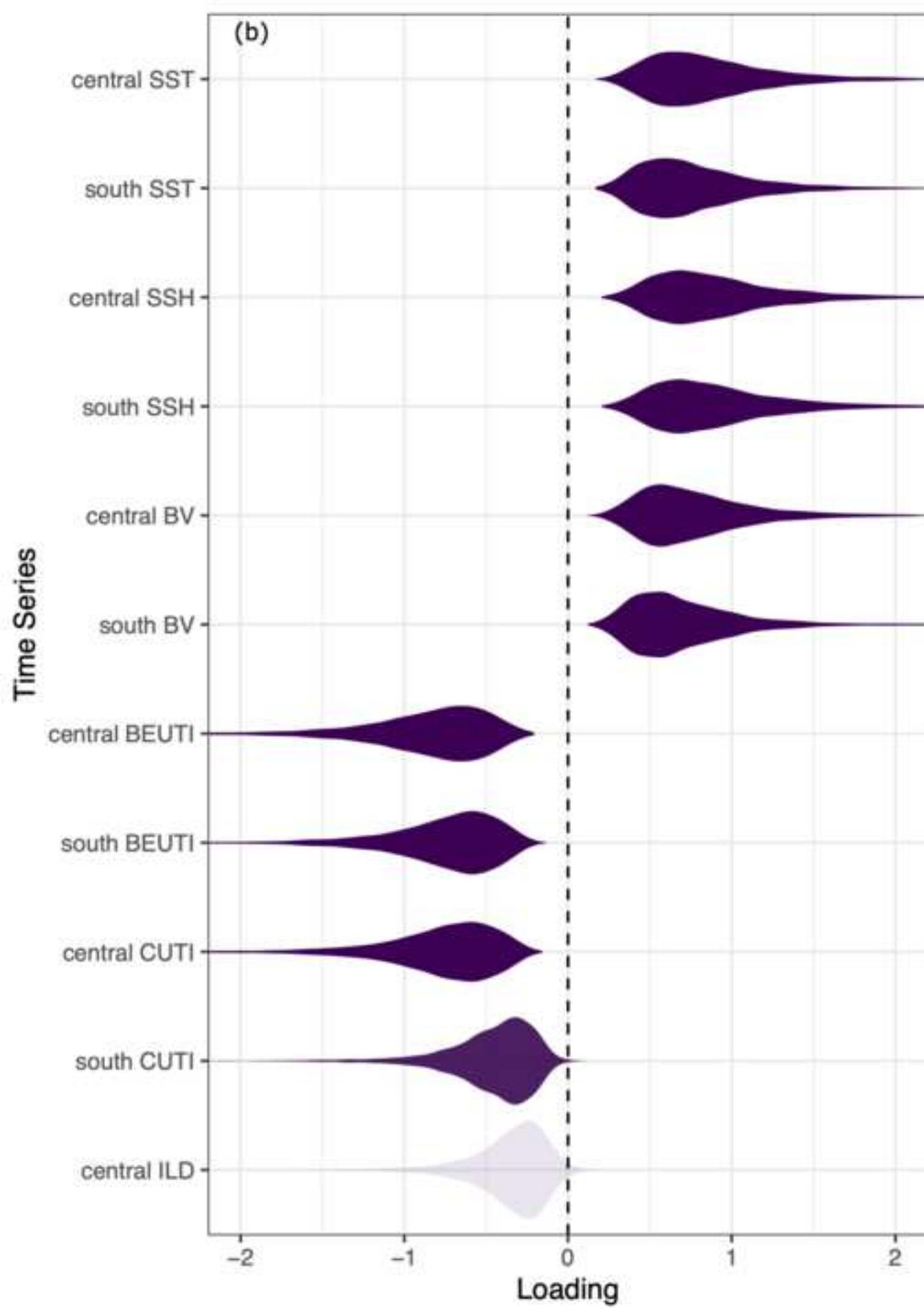
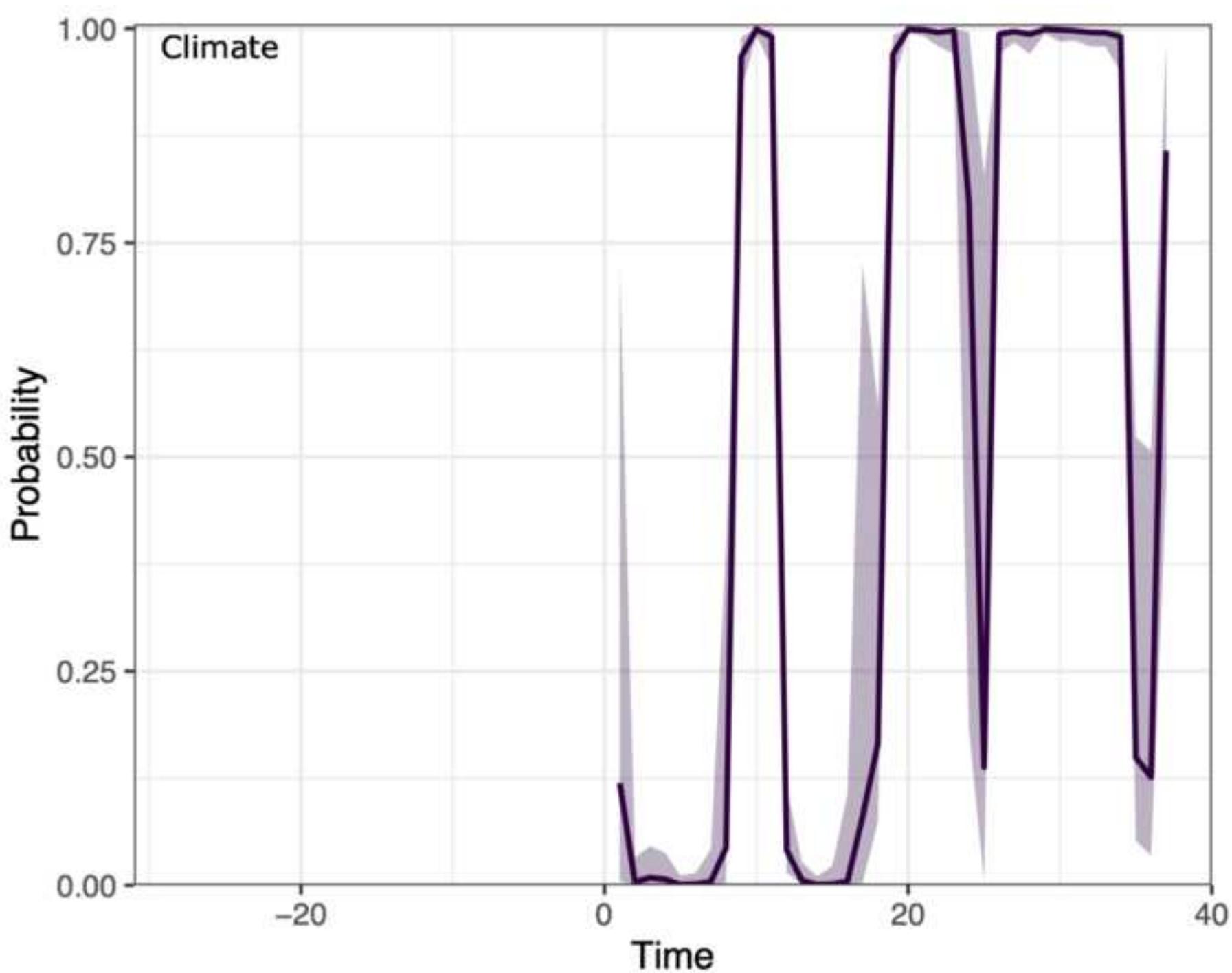
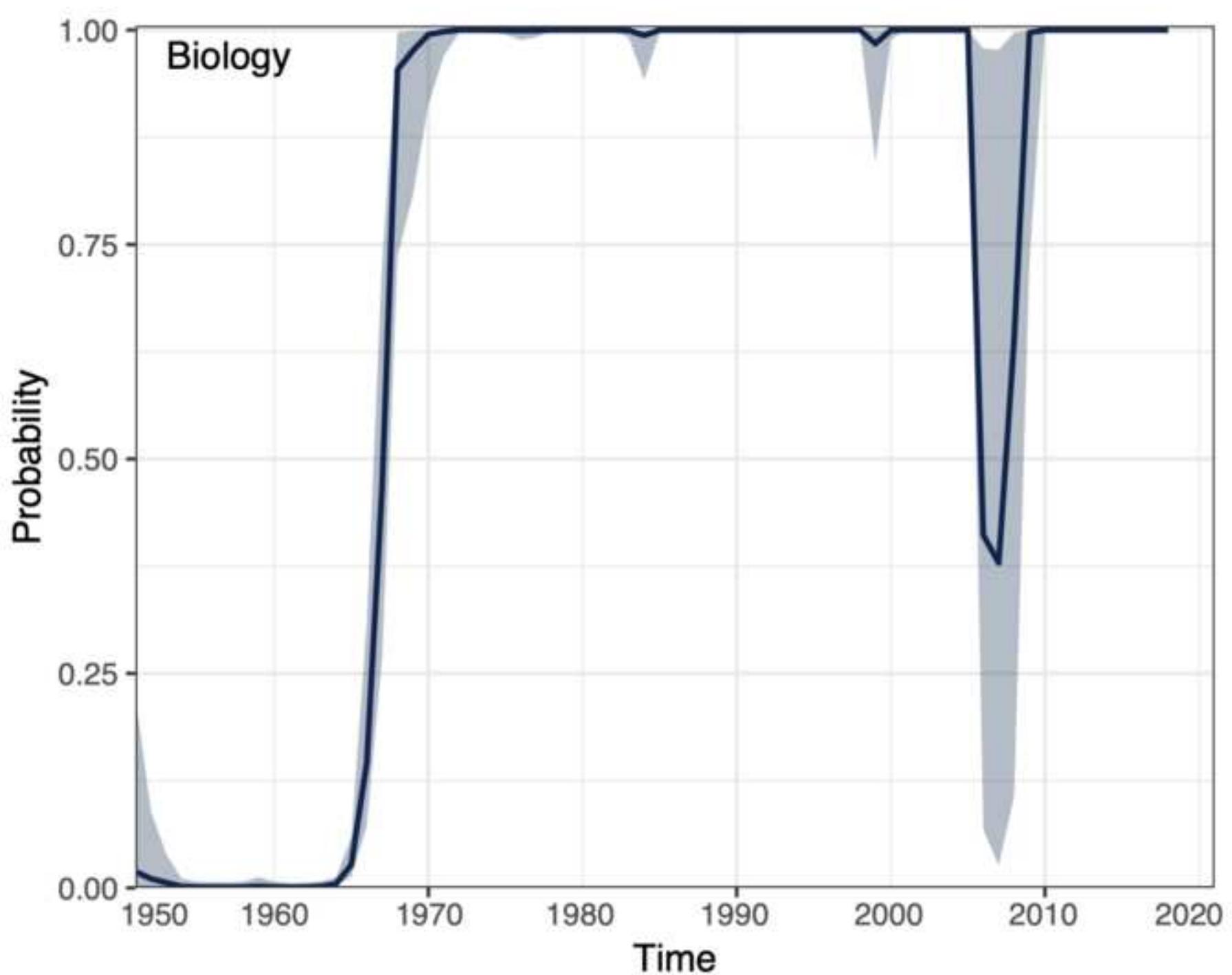
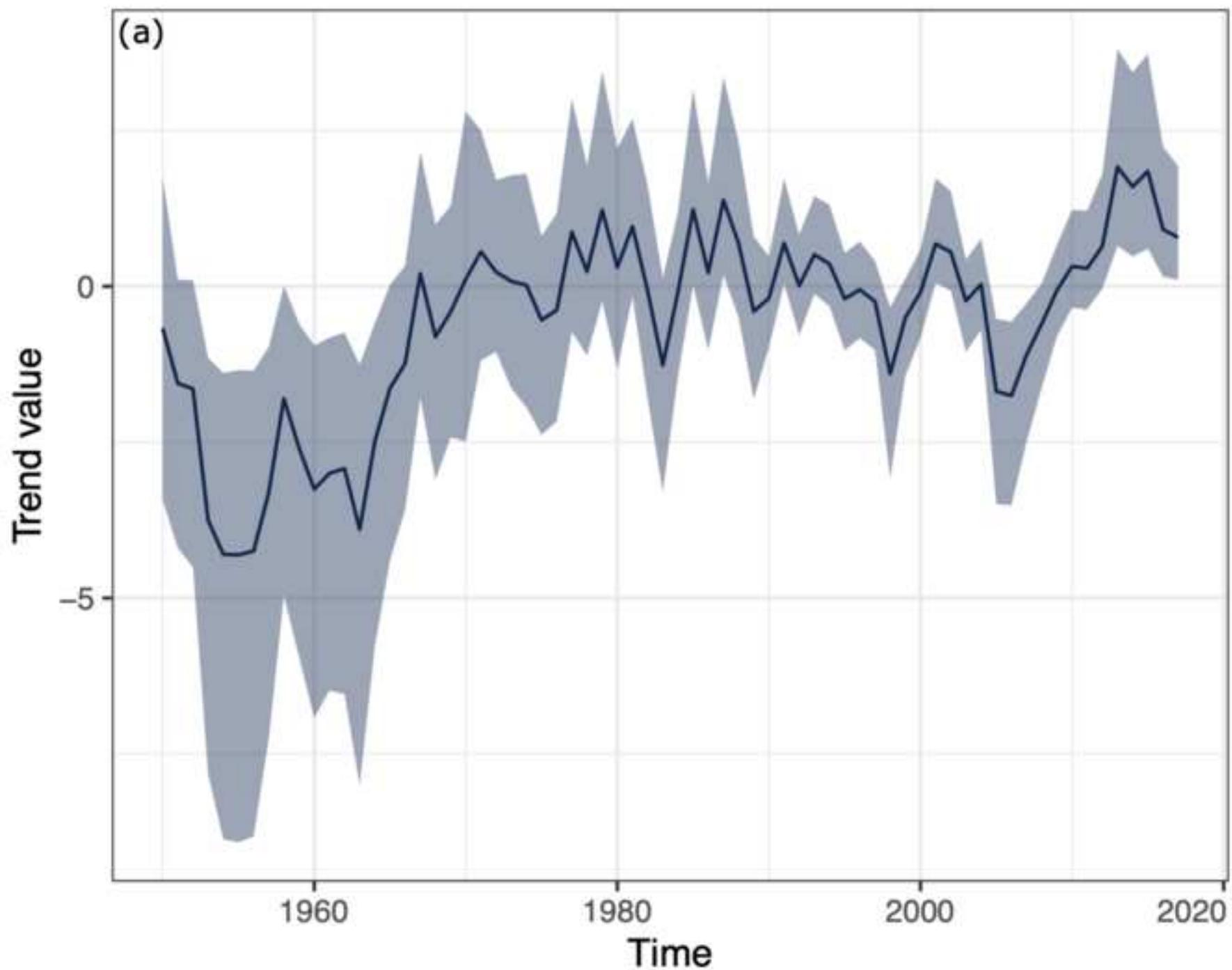


Figure 2b

[Click here to access/download;Figure;Figure_2b.tif](#)







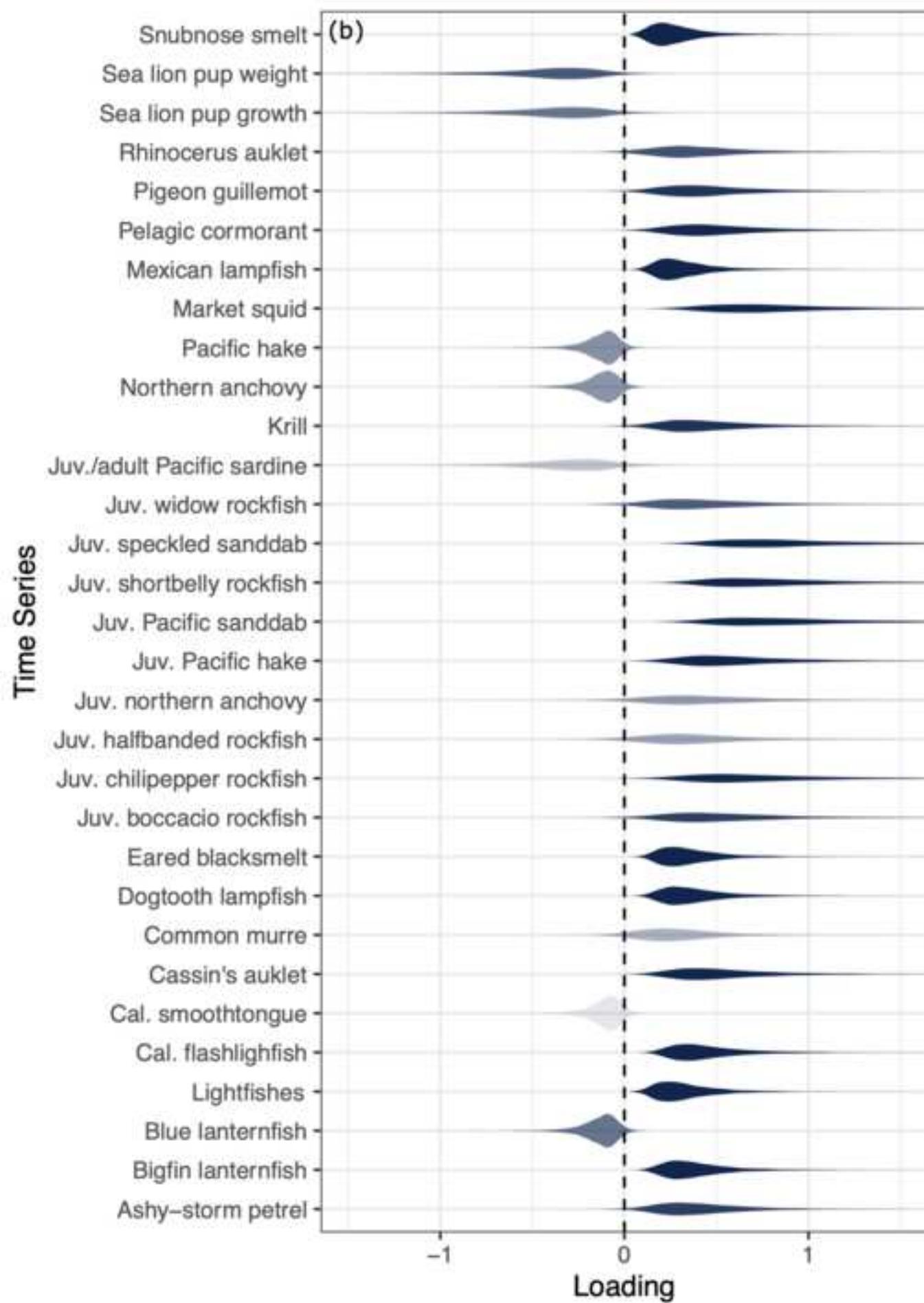


Figure 5

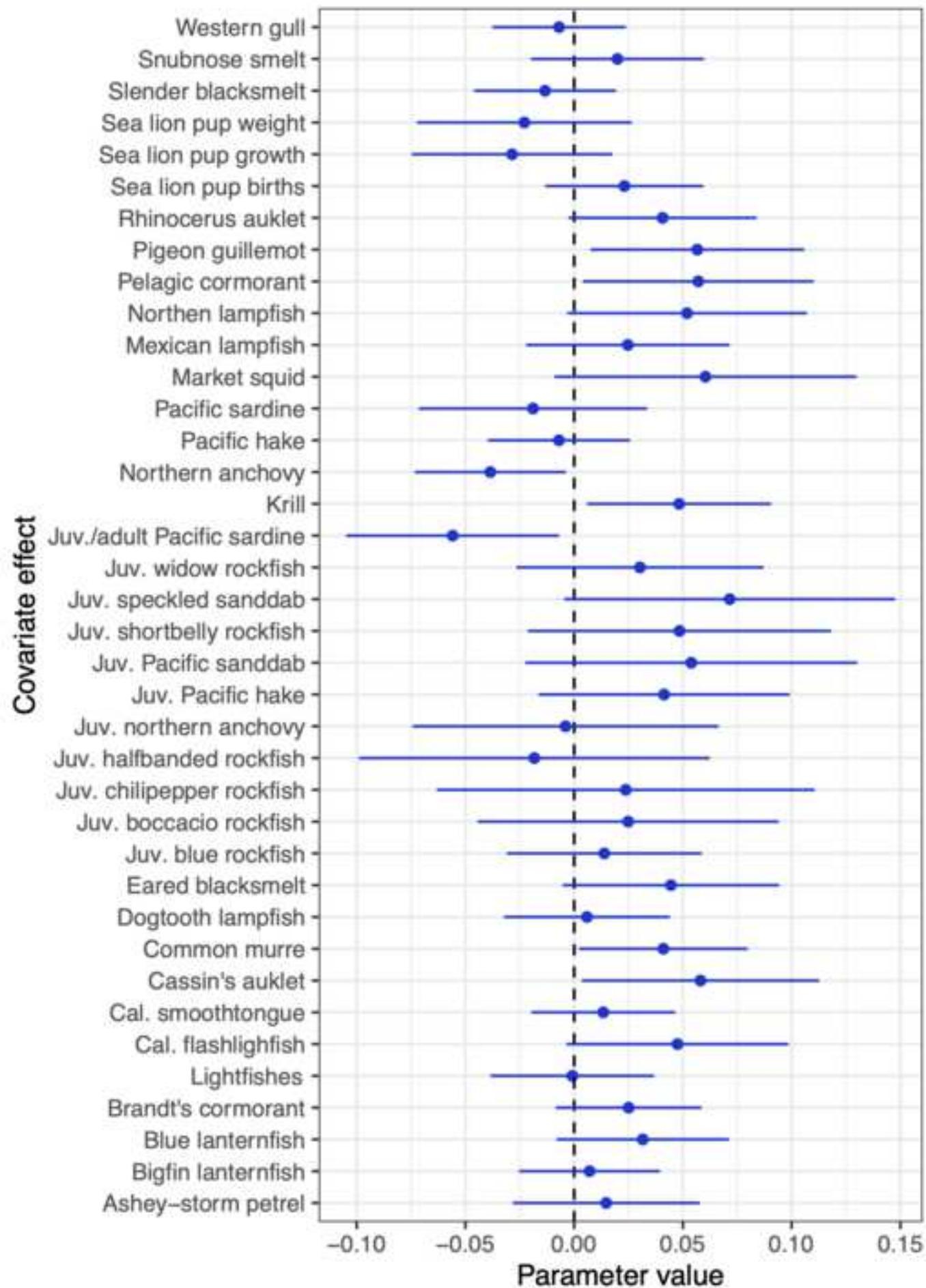
[Click here to access/download;Figure;Figure_5.tif](#)

Figure 6

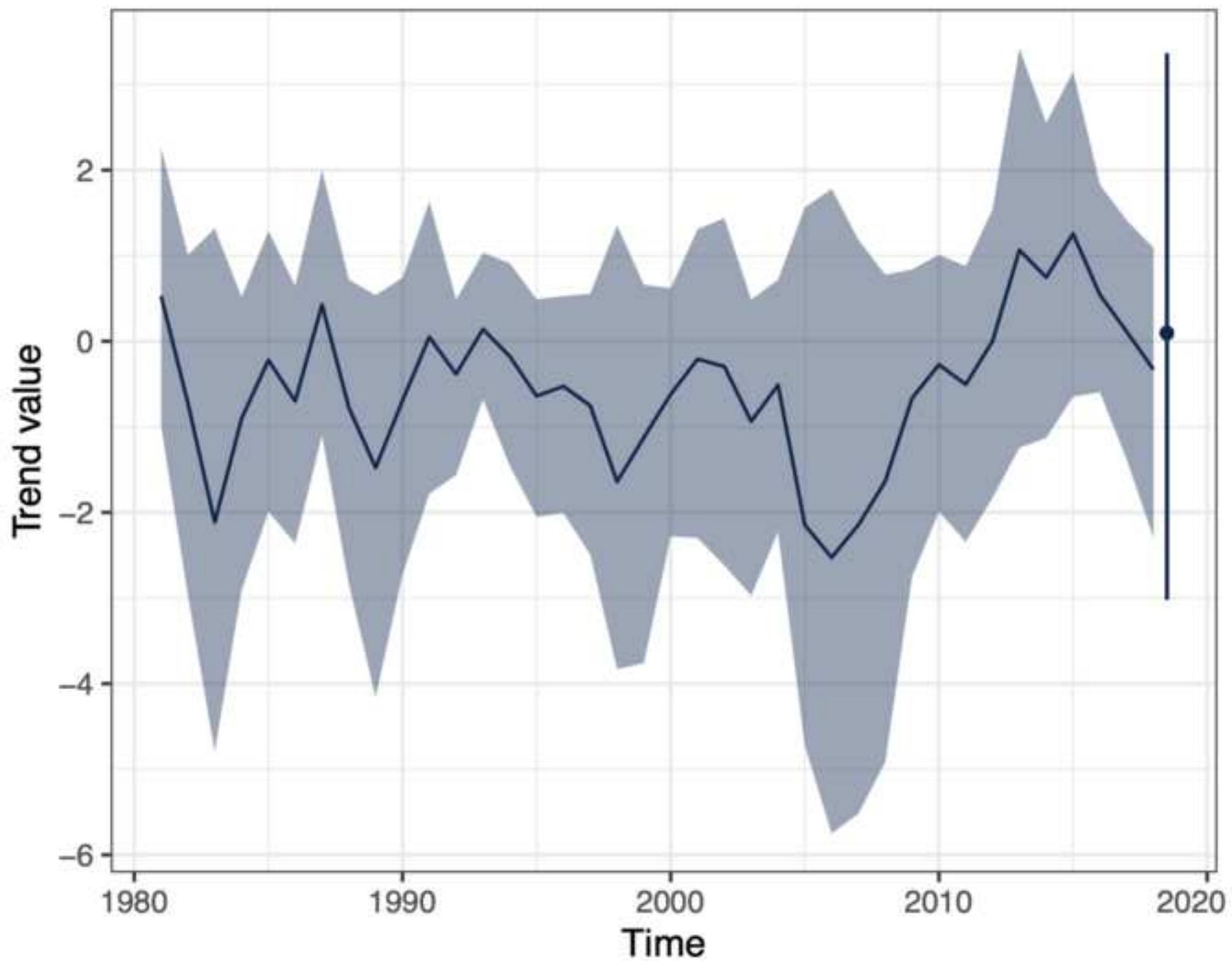
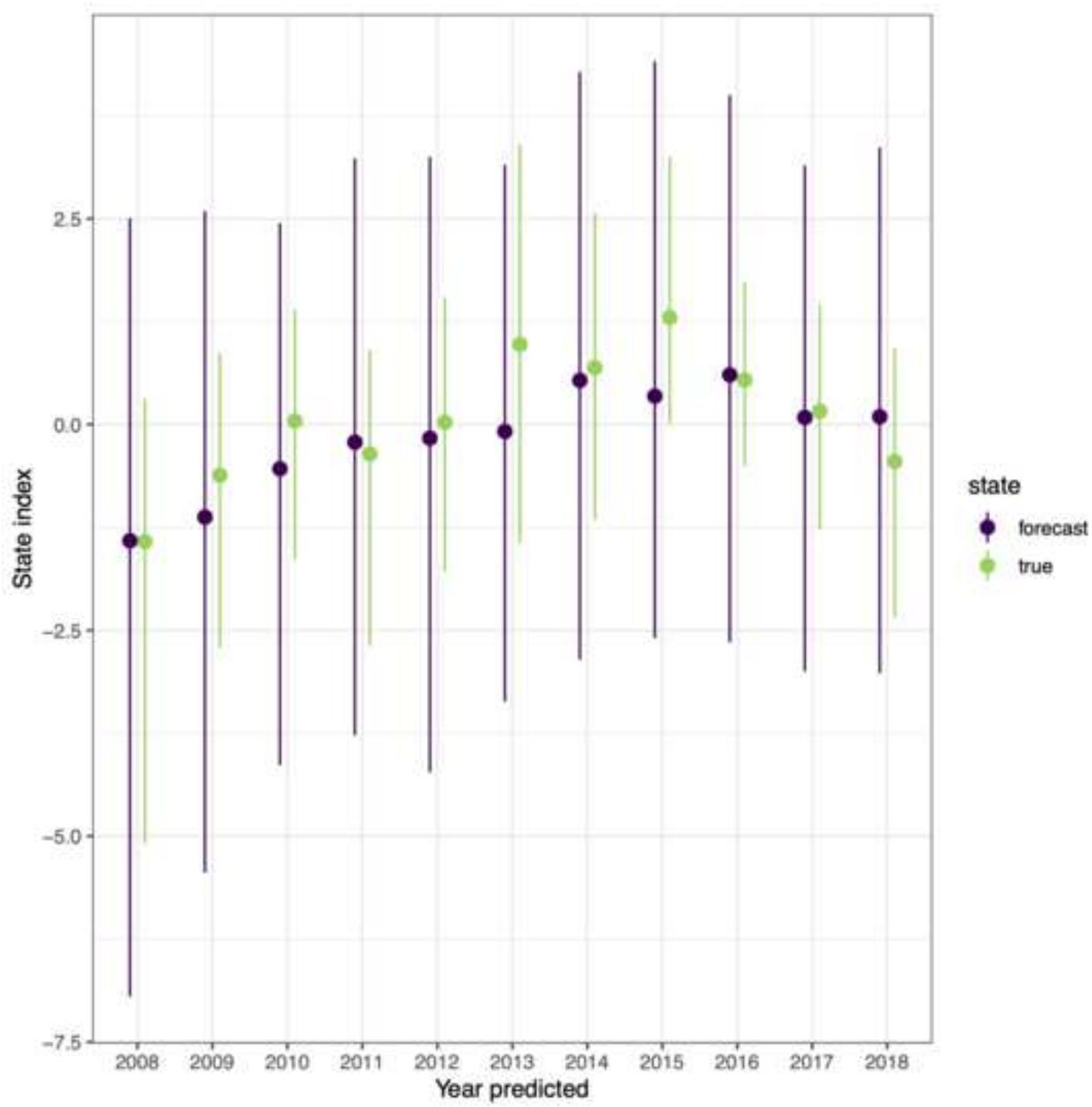
[Click here to access/download;Figure;Figure_6.tiff](#)

Figure 7

[Click here to access/download;Figure;Figure_7.tiff](#)



Click here to access/download

Supporting Information

[Hunsicker_EtAl_SupportingInfo_revised.docx](#)

1 Short Title: Tracking and forecasting community state

2

3 **Tracking and forecasting community responses to climate perturbations in the California**

4 **Current Ecosystem**

5

6 Mary E. Hunsicker^{1,†,*}, Eric J. Ward^{2,†}, Michael A. Litzow³, Sean C. Anderson⁴, Chris J.

7 Harvey², John C. Field⁵, Jin Gao⁶, Michael G. Jacox⁷, Sharon Melin⁸, Andrew R. Thompson⁹,

8 Pete Warzybok¹⁰

9

10

11 ¹Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National
12 Oceanic and Atmospheric Administration, 2032 SE OSU Drive, Newport, OR 97365 US

13 ²Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service,
14 National Oceanic and Atmospheric Administration, Seattle, WA 98112 USA

15 ³Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric
16 Administration, Kodiak, AK 99615 USA

17 ⁴Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, BC V9T 6N7 Canada

18 ⁵Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service,
19 National Oceanic and Atmospheric Administration, Santa Cruz, CA USA

20 ⁶Centre for Fisheries Ecosystems Research, Memorial University of Newfoundland, St. John's, NL A1C
21 5R3 Canada

22 ⁷Environmental Research Division, Southwest Fisheries Science Center, National Marine Fisheries
23 Service, National Oceanic and Atmospheric Administration, Monterey, CA 93940 USA / Physical
24 Sciences Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305 USA.

25 ⁸Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service,
26 National Oceanic and Atmospheric Administration, Seattle, WA 98115 US

27 ⁹Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and
28 Atmospheric Administration, La Jolla, CA USA

29 ¹⁰Point Blue Conservation Science, Petaluma, CA USA

30

31 *Corresponding author
32 Email: mary.hunsicker@noaa.gov

33

34 [†] These authors contributed equally to this work

36 **Abstract**

37 Ocean ecosystems are vulnerable to climate-driven perturbations, which are increasing in
38 frequency and can have profound effects on marine social-ecological systems. Thus, there is an
39 urgency to develop tools that can detect the response of ecosystem components to these
40 perturbations as early as possible. We used Bayesian Dynamic Factor Analysis (DFA) to develop
41 a community state indicator for the California Current Ecosystem (CCE) to track the system's
42 response to climate perturbations, and to forecast future changes in community state. Our key
43 objectives were to (1) summarize environmental and biological variability in the southern and
44 central regions of the CCE during a recent and unprecedented marine heatwave in the northeast
45 Pacific Ocean (2014–2016) and compare these patterns to past variability, (2) examine whether
46 there is evidence of a shift in the community to a new state in response to the heatwave, (3)
47 identify relationships between community variability and climate variables; and (4) test our
48 ability to create one-year ahead forecasts of individual species responses and the broader
49 community response based on ocean conditions. Our analysis detected a clear community
50 response to the marine heatwave, although it did not exceed normal variability over the past six
51 decades (1951–2017), and we did not find evidence of a shift to a new community state. We
52 found that nitrate flux through the base of the mixed layer exhibited the strongest relationship
53 with species and community-level responses. Furthermore, we demonstrated skill in creating
54 forecasts of species responses and community state based on estimates of nitrate flux. Our
55 indicator and forecasts of community state show promise as tools for informing ecosystem-based
56 and climate-ready fisheries management in the CCE. Our modeling framework is also widely
57 applicable to other ecosystems where scientists and managers are faced with the challenge of
58 managing and protecting living marine resources in a rapidly changing climate.

59 **Introduction**

60 Climate perturbations can have strong impacts on ocean ecosystems that in turn affect social and
61 economic components of human communities. These effects may be exacerbated when changes
62 in ocean conditions are more extreme, such as during marine heatwaves ([prolonged events of](#)
63 [anomalously warm ocean waters](#)). The increasing attention on these extreme events and their
64 impacts (e.g., Hobday et al. 2018, Sen Gupta et al. 2020) has invigorated a push for tools that can
65 track and detect as early as possible the response of marine communities to climate-driven
66 perturbations. Early detection, and moreover, near-term forecasts of community shifts could help
67 scientists, managers, and stakeholders better prepare for and respond to the potential
68 consequences of such shifts.

69 Climate-driven shifts in community structure tend to involve rapid change across
70 multiple populations that result in switches between contrasting community assemblages that
71 may then persist for decades. A growing number of studies have documented community
72 reorganizations in response to climate drivers (e.g., Beaugrand et al. 2008, 2015, Möllman and
73 Diekmann 2012, Wernberg et al. 2016, Peabody et al. 2018). One of the best-known examples is
74 the widespread northeast Pacific community reorganization that followed the 1976/1977 shift in
75 the Pacific Decadal Oscillation ([from a cold to warm regime](#)) (Benson and Trites 2002; Hare and
76 Mantua 2000). The abrupt change from a cool to warm ocean regime had dramatic implications
77 on ecosystem functioning and living marine resources (LMRs) throughout the region (Mantua et
78 al. 1997, Anderson and Piatt 1999; Litzow and Ciannelli 2007, Peabody et al. 2018). Since then,
79 northeast Pacific marine ecosystems have experienced several interannual or decadal
80 perturbations that do not appear to have resulted in community-wide shifts of similar magnitude.

81 However, between 2014 and 2016 these ecosystems experienced a marine heatwave that
82 involved the warmest sea surface temperature (SST) and heat content anomalies that had ever
83 been observed over large areas of the North Pacific, with SST anomalies over 6°C (Bond et al.
84 2015; Walsh et al. 2018). It was one of the most extreme heatwaves globally in its combined
85 magnitude, spatial scale, and duration (Hobday et al. 2018, Sen Gupta et al. 2020), and the
86 intense, persistent warming has been attributed to a combination of natural and anthropogenic
87 forcing (Jacox et al. 2018a; Laufkötter et al. 2020).

Formatted: Superscript

88 Several studies have documented myriad biological responses to this event. For example,
89 within the California Current Ecosystem (CCE), there were mass strandings of marine mammals
90 (Cavole et al. 2016), increased whale entanglements due to shifting prey sources (Santora et al.
91 2020), mass mortality events for marine seabirds (Cavole et al. 2016, Jones et al. 2018, Piatt et
92 al. 2020), a record-breaking domoic acid outbreak (McCabe et al. 2016), shifts in pelagic
93 macronekton and micronekton communities and species richness (Santora et al. 2017, Brodeur et
94 al. 2019, Nielsen et al. 2020), irruptions of previously rare fishes and invertebrates throughout
95 the California Current (Sakuma et al. 2016, Morgan et al. 2019, Sanford et al. 2019, Walker et al.
96 2020), and extraordinarily high recruitment of rockfishes (genus *Sebastodes*; Schroeder et al. 2018,
97 Field et al. 2021) and northern anchovy (*Engraulis mordax*; Thompson et al. 2019). Yet, to date,
98 there have been few quantitative studies of how the marine heatwave impacted the broader CCE
99 community at multiple trophic levels, and therefore the importance of this extreme event for
100 community-wide patterns of variability, and the persistence of the community response, remains
101 largely unknown.

102 Indicators of community or ecosystem state are valuable tools for tracking climate-related
103 changes in ecosystem functioning and evaluating those changes within the context of past

104 climate perturbations (Harvey et al. 2020). Moreover, combining long-term monitoring surveys
105 and data with modeling frameworks that summarize information across taxa and life stages that
106 respond quickly to climate perturbations could provide early detection of an ecosystem shifting
107 into a novel state. Early detection of such shifts would benefit ecosystem-based and climate-
108 ready fisheries management strategies aimed at mitigating possible deleterious ecological and
109 socio-economic outcomes. There is also a pressing need for forecasts of future ecosystem states
110 to support forward-looking management of LMRs (Hobday et al. 2016, Tommasi et al. 2017,
111 Jacox et al. 2020), including assessments of risk. As climate models and forecasts of ocean
112 conditions continue to improve, the time is ripe for there are burgeoning opportunities to
113 developing and testing methods that could provide near-term forecasts of community state in
114 relation to ocean conditions.

115 A challenge in summarizing ecosystem responses to perturbations is that time series used
116 to characterize the ecosystem often involve tens to hundreds of variables (species or climate
117 indices); there is often some degree of asynchrony among time series (unevenly or irregularly
118 spaced), and further, each is corrupted by the presence of observation errors. Disentangling these
119 sources of error and separating the signal from the noise is statistically challenging.
120 Traditionally, tools such as Principal Components Analysis (PCA) or nonmetric
121 multidimensional scaling have often been used for identifying leading patterns of variability in
122 multivariate datasets (e.g., Koslow et al. 2002, 2013); however, these approaches are ill-suited to
123 the analysis of time series data that are autocorrelated or non-stationary (Planque and Arneberg
124 2018). An alternative approach, Dynamic Factor Analysis (DFA), is better suited for identifying
125 shared trends that can be used as a community state indicator. DFA is specifically designed for
126 time series ordination, and avoids many of the problems associated with other multi-variate

127 approaches (Zuur et al. 2003). When applied to a collection of multivariate time series, inference
128 in DFA models focuses on estimating a smaller number of temporal patterns ('trends') that best
129 capture the variation observed. The observed data are then treated as a mixture of these trends
130 (Ward et al. 2019). Ward et al. (2019) recently developed a Bayesian implementation of DFA
131 that models shared trends offers added flexibility in model aspects over conventional approaches;
132 examples include allowing for extreme, detects "black swan" events (rare and difficult to predict
133 events; Anderson et al. 2017), and trend processes that do not follow a random walk. Output
134 from these Bayesian DFA models can also be used to and estimates the probability of extreme
135 events occurring or switches among contrasting system states. In the first application of this new
136 method, Litzow et al. (2020a) examined shared trends of climate and biology time series in the
137 Gulf of Alaska. Their study did not detect evidence for wholesale community reorganization
138 during the recent northeast Pacific marine heatwave; however, their findings indicated potential
139 for new patterns of ecosystem functioning with continued warming of ocean temperatures.

140 Here The goal of our study is to we build on this set of novel statistical tools to develop a
141 model of the CCE state that can both track and forecast ecosystem changes in response to climate
142 perturbations. More specifically, we expand the Bayesian implementation of DFA to test the
143 community response to environmental variables within the modeling framework and to develop
144 near-term forecasts of future community states. Using climate and biological data from the
145 central and southern regions of the CCE, our goals-specific objectives were to: (1) summarize
146 environmental and biological variability during 2014–2016 marine heatwave and compare these
147 patterns to past variability; (2) examine whether there is evidence assess the probability of of
148 observed departures from previous climate patterns and of switches to a new states in community
149 variabilitycommunity state during the heatwave; (3) identify relationships, if any, between

150 community variability and climate variables; and (4) test our ability to create one-year ahead
151 **simultaneous**-forecasts of species responses and the community state based on environmental
152 information. While the focus of our study is the CCE, the approach applied here is widely
153 applicable to the myriad marine ecosystems worldwide that are vulnerable to a rapidly changing
154 climate.

155

156 **Methods**

157 *Data*

158 In our analysis, we used oceanographic time series from the southern (n=6) and
159 central (n=6) regions of the CCE, derived from a data assimilative configuration of the Regional ← Formatted: Normal
160 Ocean Modeling System (ROMS) with 0.1° (~10 km) horizontal resolution and 42 terrain-
161 following vertical levels (Neveu et al. 2016; oceanmodeling.ucsc.edu). From the ROMS output,
162 we generated monthly time series covering 1980-2018 for a suite of variables including sea
163 surface temperature (SST), sea surface height (SSH), isothermal layer depth (ILD), Brunt-
164 Väisälä frequency (BV), a coastal upwelling transport index (CUTI), and a biologically effective
165 upwelling transport index (BEUTI). The ILD is similar to mixed layer depth and defines the
166 depth where temperature deviates by 0.5°C from the surface value. BV is a measure of water
167 column stratification, averaged over the upper 200 m of the water column. CUTI and BEUTI are
168 upwelling indices that quantify vertical transport and nitrate flux through the base of the mixed
169 layer, respectively (Jacox et al. 2018b). The data were annually averaged (July-June) from the
170 coast to 100 km offshore, with the exception of CUTI and BEUTI, which capture coastal
171 upwelling within 75 km of shore. In the alongshore direction, we calculated averages for two
172 regions with a division at Point Conception, California, separating the southern portion of the

173 CCE (31–34.5°N) from the central region (34.5–40.5°N, Fig. 1). This is in response to the
174 recognition of Point Conception as a major biogeographic boundary for the California Current
175 System, with differing wind and current patterns north and south of that feature (Checkley and
176 Barth 2009, Gottscho 2016). The annual averages were taken from July to June to capture the
177 influence of the El Niño–Southern Oscillation (ENSO), which peaks in winter and is the
178 dominant mode of interannual variability influencing the California Current (Jacox et al. 2015).
179 We developed models using ROMS output rather than empirical measurements because they
180 provide full spatial and temporal coverage of surface and subsurface conditions, incorporate
181 available observations, and will enable the use of ROMS forecasts to then forecast biological
182 changes in the CCE. This ocean model is constrained by available satellite and *in situ*
183 observations to improve its fidelity to nature and has been validated against independent *in situ*
184 observations (Neveu et al. 2016, Schroeder et al. 2014). Output from this model has been widely
185 used to characterize CCE oceanography, its relation to large scale climate variability, and its
186 influence over the marine ecosystem from phytoplankton to top predators (see Discussion). More
187 details on the oceanographic time series can be found in S1 Table and S1 Figure.
188

189 **Figure 1.** Sampling locations of California Current Ecosystem biology included in the study
190 analyses. Abundance data for pelagic juvenile groundfishes and invertebrates are collected on the
191 Rockfish Recruitment and Ecosystem Assessment Survey (RREAS). Ichthyoplankton data are
192 collected on the California Cooperative Oceanic Fisheries Investigations (CalCOFI) survey.
193 Seabird reproductive success and California sea lion (*Zalophus californianus*) pup time series are
194 collected on Southeast Farallon Island and San Miguel Island, respectively. See S1 Table and S1
195 Figure for detailed information on the individual time series. The base map layer was sourced

Formatted: Font: Times New Roman, 12 pt

Formatted: Font: Not Italic

Formatted: Font: Italic

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

196 from NOAA National Geophysical Data Center (2009) ETOPO1 1 Arc-Minute Global Relief
197 Model. NOAA National Centers for Environmental Information (accessed: 19 April 2013,
198 Amante, C & BW Eakins 2009).
199

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

200 The biology time series included in our analysis were selected based on three criteria: first, the
201 measured variables would be expected to show rapid (0- to 1-year lag) responses to climate
202 variability; second, the time series could be updated with no more than one year lag for
203 processing time to increase the speed at which biological responses to perturbation could be
204 detected; and third, the time series were at least 15 years long. A threshold of 15 years allowed
205 us to include long time series that spanned as many climate perturbations as possible and also
206 have enough biological time series to develop an informative indicator of community state. In
207 addition, 15 years is a threshold that has been previously used to define "long oceanographic
208 time series" in the California Current (McClatchie et al. 2014). The biology time series that met
209 our selection criteria (n=38) included ichthyoplankton, pelagic young-of-the-year (juvenile
210 fish), squid, and krill abundance; seabird productivity; and California sea lion pup body
211 condition metrics (Fig. 1, S1 Table). These 38 time series were collected from four disparate
212 ocean surveys, and span between 22 and 68 years. Datasets collected from surveys that included
213 spatial attributes (e.g., ichthyoplankton and pelagic juvenile fish surveys) were first standardized
214 using Generalized Additive Models to create a univariate time series for each species. While
215 these datasets generally include spatial random sampling, the index standardization accounts for
216 uneven distributions of effort (in space or time). Details on the standardization of individual
217 datasets are included in S1 Appendix. In addition, the biology data were normalized with log
218 transformations where appropriate (all zeros were changed to NAs). For example, if the time

Formatted: Font: (Default) Times New Roman

Formatted: Normal (Web), Indent: First line: 0"

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

219 series data were assumed to be lognormally distributed (e.g., weight/count data) or the
220 coefficient of variation was > 1, the data were log transformed. All of the time series from an
221 individual dataset (survey) were treated the same, i.e., logged or not. More details on the biology
222 time series used in this study and the associated data sources and log transformations are
223 summarized in S1 Table and S1 Figure.

Formatted: Font: (Default) Times New Roman

225 *Modeling*

226 We describe the methods in detail below, but in summary our work flow was to (1) apply
227 Bayesian DFA to climate and biology datasets separately and use model selection tools to
228 identify the best supported model and number of shared trends, (2) apply ‘black swan’ and
229 regime detection methods to detect extreme events and alternating community states,
230 respectively, (3) identify whether the CCE community state was strongly correlated with the
231 climate time series (compare performance of the biology models with/without environmental
232 covariates), and (4) evaluate our skill at making predictions of community state and individual
233 species variables. [These four steps map on to the four study objectives outlined in the](#)
234 [introduction.](#)

235 236 *Dynamic Factor Analysis*

237 We used a Bayesian version of Dynamic Factor Analysis (DFA, Zuur et al. 2003, Ward et
238 al. 2019) using the software Stan and R (R Core Team 2018) as implemented in the ‘bayesdfa’
239 package (Ward et al. 2020). DFA is a multivariate statistical tool somewhat analogous to
240 Principal Components Analyses, but for time-series data (Holmes et al. 2018, <https://cran.r-project.org/web/packages/MARSS/vignettes/UserGuide.pdf>). For a collection of time series, the

242 number of estimated ‘trends’ is specified *a priori*, and DFA estimates these latent trends as
243 independent random walks. In mathematical form, this is expressed as

244
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_{t-1},$$

245 where \mathbf{x}_t represents the value of latent (unobserved) trends at time t , and the process error
246 deviations \mathbf{w}_{t-1} are generally assumed to be white noise having arisen from a multivariate
247 normal distribution (with an identity covariance matrix for identifiability). The latent trends are
248 mapped to the observed data through an estimated loadings matrix \mathbf{Z} and residual error \mathbf{e}_t ,

249
$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{b} \cdot \mathbf{d}_t + \mathbf{e}_t,$$

250 where \mathbf{y}_t is the vector of observed states at time t , and the residual error terms \mathbf{e}_t are assumed to
251 be drawn from a univariate or multivariate normal distribution. Though the covariance matrix of
252 \mathbf{w}_t is generally fixed (Zuur et al. 2003), the covariance matrix of \mathbf{e}_t can be structured; variances
253 may be shared or not across time series, and off diagonal elements may be estimated. The
254 parameter vector \mathbf{b} represents optional estimated coefficients relating covariates \mathbf{d}_t to the
255 observed response. In the context of our DFA modeling, we included climate variables as \mathbf{d}_t in
256 models where the biological observations were used as the response \mathbf{y}_t .

257 Because we implemented the DFA model in a Bayesian setting, we were able to extend
258 this model to include additional features. First, to include extreme events, we relaxed the
259 assumption about process errors \mathbf{w}_t being drawn from a normal distribution and used a
260 multivariate Student-t distribution ([MVT](#)) instead (Anderson and Ward 2019). We also modified
261 the process equation to consider an optional vector of AR(1) coefficients $\boldsymbol{\phi}$ on the latent trends.
262 $\mathbf{x}_t = \boldsymbol{\phi}\mathbf{x}_{t-1} + \mathbf{w}_{t-1}$ (Ward et al. 2019). A final modification of the conventional DFA model is
263 that for some models, process variances can be estimated rather than fixed at 1 (maximum
264 likelihood approaches generally use this constraint for identifiability). As implemented in Stan

265 (Stan Development Team 2016, Hoffman and Gelman 2014, Carpenter et al. 2017), we
266 conducted estimation with three chains, with a warm-up of 2000 samples, followed by 2000
267 iterations. We used the split-chain potential scale reduction factor (Gelman and Rubin 1992,
268 Gelman et al. 2013) to assess convergence ($Rhat < 1.05$). Code to replicate these analyses is
269 deployed as an R package on CRAN ('bayesdfa', Ward et al. 2020) and our public Github
270 repository, <https://github.com/fate-ewi/bayesdfa>.

271

272 *Models structure optimization*

273 We ran the DFA on climate datasets (1981–2017) and biological datasets (1951–2017)
274 across the southern and central regions of the California Current combined. Running the analysis
275 at this spatial scale allowed us to capture the broader community response to climate
276 perturbations, compared to running models on each multivariate dataset independently (e.g., time
277 series from a single survey). There are a number of ways to evaluate predictive accuracy of these
278 models. [The commonly used LLeave-One-Out Cross-Validation \(LOO-CV\)](#), for example holds
279 each observation out in turn and predictions are made from the remaining data. As our focus was
280 on the temporal nature of the data and forecasting component, we implemented a variant of k-
281 fold cross validation and treated individual years as unique ‘folds’. Because our objectives
282 involved evaluating these models for future predictions, we implemented the Leave-Future-Out
283 Cross Validation Information Criterion (LFO-CV, Bürkner et al. 2020). We used this approach to
284 identify data support for (1) the number of latent DFA trends ($n = 1-3$), (2) first-order
285 autoregressive AR(1) coefficients on the trends (ϕ estimated with a $Normal(0,1)$ prior), (3)
286 Student-t deviations (i.e., evidence of extreme events, using a prior on the MVT degrees of

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

287 freedom parameter, ν , of $\nu \sim \text{Gamma}(2, 0.1)$), and (4) a fixed versus estimated trend variance
288 (using a prior on the standard deviation, σ_w , of $\sigma_w \sim \text{Normal}(0,1)$).

289 In addition, we used LFO-CV to identify the most appropriate error structure for the
290 climate dataset—specifically whether the times series had equal (shared) or unequal (unique)
291 observation errors. For the biology models, we assumed the observation errors were unique by
292 dataset, and our estimates of survey variance supported this assumption.

293 For each model formulation, we applied the LFO-CV method by first fitting the model to
294 all years of data prior to year T (i.e., training data, years 1, 2, ..., ($T-1$)) and then using the fitted
295 model to predict the trend value in year T (i.e., test data). We repeated this process for 10 years,
296 starting with 2017 as year T and working back to 2008, and then calculated the expected log
297 predictive density (ELPD) across those time steps. The climate and biology models with the
298 highest ELPD were deemed the best supported models. The LFO-CV is a preferred method for
299 evaluating future predictive performance of Bayesian models because it properly accounts for
300 time series structure, and unlike other Bayesian cross-validation methods, does not produce
301 overly optimistic estimates (Bürkner et al. 2020).

302

303 *Detection of extreme events and regime shifts*

304 After identifying the best-supported DFA model for the climate and biological datasets,
305 we conducted a post-hoc examination of outlier detection and regime shifts. For outlier detection
306 of black swan events, we implemented a method similar to that described in Anderson et al.
307 (2017) and applied it to the climate and biology time series. This approach relies on first
308 differencing the posterior trend mean estimates of the climate and biology trends, $x_t - x_{t-1}$ and
309 then applies a normal density function to identify year-over-year changes that were unlikely to

Formatted: Font: Italic

310 have arisen from a normal distribution (given the process variance). Probabilities can then be
311 assigned to the deviations in each year (e.g., ‘there is a 1:1000 chance of observing a deviation
312 similar to that estimated in year t ’). As described in Ward et al. (2019), the presence of regimes
313 can also be estimated by applying hidden Markov models (HMM) to the estimated state indices
314 from a DFA. We evaluated support for regimes and alternate states by using the posterior trend
315 estimates from each model as input. The Bayesian Leave-One-Out Cross Validation
316 Information Criterion (LOO-CV, Vehtari et al. 2017) was used to identify the data support for
317 the number of trends-regimes ($n = 1–3$). The model with the lowest LOO-CV value is deemed
318 the best model.

319

320 Climate-biology relationships and forecasts of community state
321 While a wide variety of multivariate or univariate time series methods could be applied to our
322 observed time series to generate forecasts, our objectives were to develop simultaneous estimates
323 of both the community state (i.e., the DFA trend value) (s) and the raw time series (i.e.,
324 individual time series summarized by the biology DFA model). We evaluated the ability of our
325 DFA models to generate short-term (one year lead-time) forecasts of community state by first
326 evaluating whether the performance of the biology DFA model was improved when climate time
327 series were included as covariates in the model. If climate time series were found to better
328 explain the variability in the biology time series, these relationships could potentially be used to
329 forecast community trends. For our analysis, we ran the DFA on a subset of the biology data
330 overlapping in time with the climate dataset, i.e., 1981–2017, to make out-of-sample predictions.
331 We used the same LFO-CV procedure described above, with the same forecast period (2009–
332 2017) to compare the biology models with and without a single climate covariate (see S2 Table

Formatted: Font: Not Italic

Formatted: Font: Not Italic

333 for all model formulations). In this case, the model used biological and climate data from all
334 preceding years and climate data from the year to be forecast. The six climate covariates from
335 the southern region and the central region of the CCE (12 total) were tested in this analysis. Once
336 the best-supported biology-covariate model was identified, we used that model to make
337 predictions of individual species parameters and the community state (i.e., DFA trend value) in
338 2018 using climate data from that same year and the raw time series of the individual species
339 (i.e., the biology time series summarized by the DFA model). We evaluated forecast skill based
340 on the prediction errors of individual species parameters individual species time series and by
341 comparing the forecasts for 20089–2018 to the 20089–2018 trend values estimated from the
342 biology-covariate model that only included data prior to the forecast year.

Commented [MEH1]: Eric – can you verify that this is correct and that we have the correct figure for this in the scc repo? Thank you!

Formatted: Font: (Default) Times New Roman, Font color: Auto, Pattern: Clear

344 Results

345 Climate and biology trends

346 The model with the highest predictive accuracy (ELPD) of the climate state in the southern and
347 central regions of the CCE was a one-trend DFA model (Model 1 in Table 1, Fig. 2a). This
348 model included unique observation variances across the six time series, support for heavy-tailed
349 deviations of the latent trend, an AR(1) coefficient on the trend (S1 Fig.), and an estimated trend
350 variance. Overall, the trend captured a well-documented cooling period in the CCE between
351 1980 and 2010 (e.g., Seo et al. 2012), as well as strong El Niño events (e.g., 1982–1983, 1997–
352 1998, 2015–2016) and the 2014–2016 marine heatwave. The trends and loadings indicate that
353 these events were generally associated with weaker upwelling, reduced mixed layer depth, low
354 nutrient flux, and warm, stratified waters (Fig. 2a, b).

355 All but one of the climate time series (central ILD) were strongly associated with the single
356 trend, i.e., at least 90% of the loading posterior distributions associated with each time series
357 were above or below zero (Fig. 2b). The SST, SSH, and BV frequency (water column
358 stratification) time series from the southern and central regions of the CCE loaded positively on
359 this trend (Fig. 2b). The BEUTI and CUTI time series from both regions of the CCE and the ILD
360 time series from the central region loaded negatively on the trend (Fig. 2b). ~~Overall, the trend~~
361 ~~captured a well documented cooling period in the CCE between 1980 and 2010 (e.g., See et al.~~
362 ~~2012), as well as strong El Niño events (e.g., 1982–1983, 1997–1998, 2015–2016) and the 2014–~~
363 ~~2016 marine heatwave. The trends and loadings indicate that these events were generally~~
364 ~~associated with weaker upwelling, reduced mixed layer depth, low nutrient flux, and warm,~~
365 ~~stratified waters (Fig. 2a, b).~~

366

367 **Table 1.** Summary information for climate and biology Bayesian DFA models, including
368 whether process error was estimated, observation error variances (unequal or equal among time
369 series, or unique to each survey), the number of model trends, expected log pointwise predictive
370 densities (ELPD), and standard error of ELPD. Bold text highlights the models that show best
371 support or highest predictive accuracy for the climate and biology data for the southern and
372 central California Current ecosystem (i.e., highest ELPD). All climate and biology models
373 include an AR(1) process and Student-t deviations.

Time series	Model	Process sigma	Variance index	Trends	ELPD	SE ELPD
Climate	1	Yes	unequal	1	-10551.89	759.44
	2	No	unequal	1	-10682.39	712.96
	3	Yes	equal	1	-16793.59	1732.54
	4	No	equal	1	-16881.03	1824.42
	5	No	unequal	2	-17441.01	1655.23
	6	Yes	unequal	2	-17818.86	1813.65
	7	Yes	unequal	3	-21571.65	1674.77
	8	No	unequal	3	-22882.99	1695.12
	9	No	equal	2	-23927.86	2694.72
	10	Yes	equal	2	-24536.19	2953.38
	11	Yes	equal	3	-38895.66	4024.93
	12	No	equal	3	-38920.87	4762.15
Biology	13	No	survey	1	-2003.32	95.76
	14	Yes	survey	1	-2003.96	88.88
	15	No	survey	2	-2155.52	76.05
	16	Yes	survey	2	-2198.79	78.86
	17	No	survey	3	-2334.55	118.31
	18	Yes	survey	3	-2335.10	112.70

374

375

376 **Figure 2.** Climate variability in southern and central California Current ecosystem: a) shared
 377 trend with 95% credible intervals (1981–2017), b) posterior distributions for loadings on all of
 378 the individual time series. Loadings with darker shading indicate time series loading most
 379 strongly on the climate trend. SST, sea surface temperature; SSH, sea surface height; ILD,
 380 isothermal layer depth; BV, Brunt-Väisälä frequency (stratification); CUTI, Coastal Upwelling
 381 Transport Index; BEUTI, Biologically Effective Upwelling Transport Index. See S1 Table and
 382 S1 Figure for climate times series details.

383

384 The climate state during the marine heatwave, as indicated by the DFA trend, was within
 385 the bounds of previous observations. While there was support in the best model for heavy-tailed
 386 deviations in the climate trend (i.e., Student-t deviations S2+ Fig.), our post-hoc examination of
 387 outliers detected a single extreme event in the climate state in mid-1998 to mid-1999
 388 (threshold=0.001), when there was a shift from strong El Niño (1997–98) to strong La Niña

389 (1998–1999) conditions, and not around the time of the heatwave. Application of the Bayesian
390 HMM to the climate trend most supported the presence of two hidden states, reflecting the
391 probability of being in a state associated with warmer conditions versus one with cooler
392 conditions (LOO-CVIC: one-state = 129.1, two-state = 9.4, three-state = 27.2, Fig. 3). The LOO-
393 CVIC did not provide support for a shift to a third novel climate state in the southern and central
394 regions of the CCE during the marine heatwave, however there is a shift back to the previously
395 observed warm state during the marine heatwave.

396

397 **Figure 3.** Results of Hidden Markov Models (HMM) showing state probability for latent trends
398 in the climate (top) and biology (bottom) data sets. The best model for both data sets invoked
399 two states, and the median probability (and 95% credible intervals) of being in one state versus
400 the other is shown. The top figure reflects the probability of being in the state associated with
401 warmer conditions versus one with cold conditions, and bottom figure indicates that ecosystem
402 did not shift into a new state following the marine heatwave.

403

404 The best model for community variability among our biological time series was also a
405 one-e-trend model (Model 13 in Table 1, Fig. 4a). The model formulation was similar to the best
406 climate model, except the observation variances were unique by dataset (survey) and not
407 individual time series. We note that the top two models (Model 13 and 14) showed similar
408 predictive accuracy ($\Delta \text{ELPD} < 1$) and only differed with respect to whether the process variance
409 was fixed at 1 or estimated. Here we only show results for the model with a fixed process
410 variance. The biology showed strong coherence in community signal; a majority of the time
411 series (31 of 38) loaded strongly (probability > 0.9) on the single trend and most of them

412 demonstrated loadings in the same direction (Fig. 4b). The magnitude and direction of the
413 estimated loadings were consistent with the observed high relative abundance of most juvenile
414 groundfishes (rockfish, flatfish), squid, krill, and some ichthyoplankton species ~~increased~~ during
415 the marine heatwave, and suggest that the reproductive success of some seabird species was
416 higher around the time of the heatwave as well. The few time series loading in the other direction
417 on the trend indicated a reduction in sea lion pup growth rate and lower abundances of
418 juvenile/adult Pacific sardine *Sardinops sagax* and some ichthyoplankton species (e.g., larval
419 northern anchovy and Pacific hake *Merluccius productus*) associated with the heatwave.

420

421 **Figure 4.** Community variability in the southern California Current ecosystem: a) shared trend
422 with 95% credible intervals (1951–2018: [marine heatwave occurred 2014–2016](#)), b) posterior
423 distributions for loadings on individual time series (only time series with $\geq 90\%$ of the loading
424 distributions above or below zero are shown). Loadings with darker shading indicate time series
425 loading most strongly on the biology trend. Cal. = California, Juv. = juvenile fish stage,
426 Juv./adult = juvenile and adult fish stages combined, all other fish are larval fish. See S1 Table
427 and S1 Figure for times series details.

428

429 The estimated trend from this biology DFA model demonstrates a potential shift in community
430 state in the mid-1960s, although there is considerable uncertainty around the trend during this
431 early portion of the time series, likely due to the limited number of observations
432 (ichthyoplankton only) pre-dating the 1970s (Fig. 4a, S1 Fig.). ~~Similarly, the limited number of~~
433 ~~biological time series and missing sampling years in the late 1960s and 1970s, and the lack of~~
434 ~~climate outputs from ROMS prior to 1980, make it difficult to detect a documented transition~~

435 that happened in the mid 1970s (S1 Fig., see discussion). The community state appears to be
436 relatively stable from the late 1970s through the early 2000s, and the trend reached a peak
437 around 2013–2015. Evidence of a community shift early in the time series is supported by our
438 regime detection analysis, which demonstrated that a two-state model best described the latent
439 trend (LOO-CVIC: one-state = 216.4, two-state = 11.8, three-state = 41.8, Fig. 3). This shift
440 coincides with a strong increase in the abundance of a few species during that period, including
441 eared blacksmelt (*Lipolagus ochotensis*), slender blacksmelt (*Bathylagus pacificus*), northern
442 lampfish (*Stenobrachius leucopsarus*), which are cool water associated mesopelagic species, as
443 well as a rise in northern anchovy (*Engraulis mordax*) abundance prior to the shift (S1 Fig.). Our
444 analysis does not document a shift to a novel community state in response to the recent marine
445 heatwave.

446 While this model provided slight support for heavy-tailed Student-t deviations in the
447 latent trend (S32 Fig.), we did not detect any black swan events in the community state. We note
448 that the community response to two strong El Niño events (1982–1983 and 1997–1998) and to
449 unusually low productivity conditions (2005) in the central CCE appear similar in magnitude and
450 duration to the response to the 2014–2016 marine heatwave, although the directions of the
451 responses were opposite (Fig. 4a). Our regime detection analysis also captured the change in the
452 central CCE community in the mid to late 2000s (Fig. 3), which may be associated with the large
453 changes in the reproductive success of multiple seabirds (e.g., Cassin's auklet
454 *Ptychoramphus aleuticus*, common murre *Uria aalge*, Brandt's cormorant *Urile penicillatus*) and
455 in sea lion pup births around that time (S1 Fig.). These taxa may have been impacted by changes
456 in the abundance or availability of important prey items resulting from unproductive ocean

457 conditions in the central CCE in 2005 and the below normal SSTs associated with the 2007-2008
458 La Niña Event (McClatchie et al. 2008, 2009, Bjorkstedt et al. 2010).
459

460 *Forecast of community state*

461 In comparing models of the biological response with and without climate covariates, we found
462 that several biology models with climate predictors outperformed the biology models that did not
463 include covariates (S2 Table). The climate covariate resulting in the best future predictions of
464 community state was BEUTI (central region), followed by CUTI (central region) (Table 2, see
465 S2 Table for all models). The coefficients linking BEUTI to observed time series ([B-matrix](#))
466 indicate strong, positive relationships between nitrate flux and the reproductive success of
467 seabirds and the abundance of krill in the central California Current (Fig. 5). They also indicate
468 strong, negative relationships between nitrate flux and the abundance of juvenile/adult Pacific
469 sardine and larval northern anchovy (Fig. 5). The remaining biology-BEUTI relationships were
470 moderate (e.g., ichthyoplankton, market squid *Doryteuthis opalescens*) to weak (e.g., rockfish
471 spp., Fig. 5). The biology-CUTI model was similar to the biology-BEUTI model with respect to
472 model structure and estimated species loadings. The estimated coefficients in the CUTI model
473 ([S4 Fig.](#)) also show a similar pattern to those in the biology-BEUTI model ([S5 Fig.](#)). The
474 remaining covariate models only showed weak climate-biology relationships (e.g., S5-S6 Fig.).

475

476 **Table 2.** Summary information for the top biology-covariate Bayesian DFA models for each
477 covariate and the top two biology only models (years 1981-2017). The table indicates whether
478 process error was estimated ('Yes') or fixed ('No'), the number of model trends, expected log
479 pointwise predictive densities (ELPD), standard error of ELPD, the environmental covariate

480 included in the model, and the region in the California Current over which the covariate was
481 aggregated. All models had an AR(1) coefficient on the trend and Student-t deviations. Model 1
482 was deemed the best model based on its highest predictive accuracy (highest ELPD value)
483 compared to all other models. BEUTI = Biologically Effective Upwelling Transport Index; BV =
484 Brunt-Väisälä frequency; CUTI = Coastal Upwelling Transport Index; ILD = Isothermal Layer
485 Depth; SST = Sea Surface Temperature; SSH = Sea Surface Height. See S2 table for the full
486 suite of model comparisons.

487

Model	Process sigma	Trends	ELPD	SE ELPD	Covariate	Region
1	No	1	-1878.71	71.52	BEUTI	central
2	No	1	-1905.27	81.98	CUTI	central
3	No	1	-1914.62	83.22	ILD	south
4	No	1	-1927.33	88.26	SST	south
5	No	1	-1928.03	88.39	BV	south
6	No	1	-1946.75	86.33	SSH	south
7	No	1	-1951.44	73.59	None	-
8	Yes	1	-2038.71	91.44	None	-

488

489

490 **Figure 5.** A summary of the effect of the Biologically Effective Upwelling Transport Index
491 (BEUTI), a measure of nitrate flux through the base of the mixed layer, on the single species
492 parameters. Cal. = California, Juv. = juvenile fish stage, Juv./adult = juvenile and adult fish
493 stages combined, all other fish are larval fish. Blue error bars reflect 95% credible intervals.
494 Figures S4-S6 show effects of other environment covariates on the biological variables.

495

496 Given that the biology-BEUTI model was the best supported model over the null model
497 (a model without covariates), we were interested in evaluating the ability of this covariate model
498 to forecast the community state. Comparisons between the community state (out-of-sample

499 estimates) and the community state forecasts indicate that we had skill in forecasting community
500 state one year in advance (Fig. 6, S7 Fig., S8 Fig.). Forecasts of the community trend values for
501 ~~nine-ten~~ additional years (200~~89~~–201~~77~~, S~~8~~ Fig. ~~7~~) also indicate that we had some skill for many
502 of the years tested. There are wide confidence intervals around the forecasts; however, given our
503 methodology we can expect that the credible intervals around the trend forecast will be larger
504 than the historical credible intervals (Fig. 6). Forecasts have more uncertainty than historical
505 values because the variance of a random walk increases linearly with time (Holmes 2004; Ward
506 et al. 2014). Furthermore, our credible intervals are increased because we are additionally (1)
507 propagating full parameter uncertainty across the MCMC draws projecting it, and (2) using a
508 Student-t distribution, which has heavy tails and therefore makes the uncertainty intervals wider
509 than if we used normal distribution.

510

511 **Figure 6.** Community variability and forecast of the community state in the southern California
512 Current. The shared biology trend (blue line, with 95% credible intervals) derived from biology-
513 BEUTI model fit to subset of data (1981–2018) is shown along with the trend forecast for 2018
514 (circle, with 95% credible intervals). See S7 Fig. for model loadings.

515

516 **Figure 7.** Forecasts and model estimates of the ‘true’ community state in the southern and
517 central California Current in years 2008–2018 (circle, with 95% credible intervals).

518

519 Overall, the model forecast skill of individual species parameters was moderate to high
520 for half of the species included in the biology-BEUTI model (S3 Table, S~~89~~ Fig.). It is important

Formatted: Font: Bold
Formatted: Line spacing: Double

Formatted: Indent: First line: 0"

522 to emphasize that the source of variability in predictions for each of the original time series is a
523 mixture of the magnitude and uncertainty around the trends and loadings (x_t , Z), and the
524 magnitude and uncertainty in the estimated covariates (b). On the one hand, the time series
525 associated with the highest predictive skill (i.e., lowest prediction errors) included seabird
526 reproductive success (Common murre *Uria aalge*, Cassin's auklet *Ptychoramphus aleuticus*)
527 and the abundance of juvenile Pacific sanddab *Citharichthys sordidus*, juvenile halfbanded
528 rockfish *Sebastodes semicinctus*, market squid, and several ichthyoplankton species (S3 Table, S⁸⁹
529 Fig.). On the other hand, forecast skill was lowest (i.e., highest prediction errors) for the
530 abundance of some juvenile rockfishes (chilipepper *Sebastodes goodei* and widow rockfish
531 *Sebastodes entomelas*) and larval fishes (northern anchovy, mesopelagics), which is likely
532 attributed to a lag or mismatch in the timing of the climate-biology relationships. These patterns
533 in forecast skill are similar to those based on the biology-CUTI, -SST, and -ILD models (S3
534 Table). Lastly, the uncertainty around model predictions of species parameters appears to be
535 driven more by the precision of the model coefficients than by the loadings on the community
536 trend (e.g., S⁹⁴⁰ Fig.).
537

538 Discussion

539 We applied a novel set of statistical tools to data from the southern and central regions of the
540 CCE to document the community response to climate perturbations over the past six decades and
541 to create near-term forecasts of community state. Our analysis detected a community response to
542 the 2014-2016 northeast Pacific marine heatwave; however, it did not exceed normal variability
543 within the study timeframe or result in a shift to a novel community state, based on the biological
544 time series investigated here. We identified relationships between community state and multiple

545 climate variables, with nitrate flux through the base of the mixed layer having the strongest
546 correspondence with individual species time series and the shared trend in community
547 variability. Moreover, we demonstrated skill in creating simultaneous one-year lead time
548 forecasts of species responses and community state.

549

550 *Long-term changes in community state*

Formatted: Font: Italic, Font color: Black

551 ~~There have been several~~Many studies and anecdotal ~~evidence~~accounts have documented
552 ~~ef~~unexpected biological responses to the 2014–2016 northeast Pacific marine heatwave. Based
553 on the biological time series included in our analysis, the broader CCE community demonstrated
554 a ~~clear~~ response to the marine heatwave ~~(Fig. 4a), i.e. the biology trend is higher over those years~~
555 ~~than in previous years~~. However, our results do not demonstrate a widespread community
556 reorganization beyond the archetypal community structure of this ~~recognizably~~ dynamic
557 ecosystem ~~within the southern and central CCE~~ in response to this event. Instead, the mean
558 values for the shared trend in the biology time series, as well as for the shared climate trend,
559 were within the range of previous observations. ~~There were M~~many species ~~were~~ present during
560 the marine heatwave that are not typically observed ~~in in sampling or associated with the CCE.~~
561 While our analysis could not include these sporadically occurring taxa, due to the large number
562 of zero observations in the historic survey data, the ,but those exceptional presences and high
563 abundances of those warm species did not result in a persistent signal among the species for
564 which we have time series contributing to included in the DFAs. As additional years of data
565 become available, the DFA models could reveal different outcomes from 2014–2016. However,
566 this is unlikely given that the taxa and life stages used in both studies are known to respond

567 quickly to changes in ocean conditions and given our assumption that the surveys are
568 consistently sampling at the right time and location to fully characterize the short-term response.

569 There were many species present during the marine heatwave that are not typically
570 observed in sampling or associated with the CCE, but those exceptional presences and high
571 abundances of warm species did not result in a persistent signal among the species for which we
572 have time series contributing to the DFA. While our study did not detect a shift in community
573 state in the southern and central CCE during the 2014–2016 heatwave, we did detect a shift in
574 the 1960s. The 1960s shift was likely due to a regime shift previously detected in the southern

575 California ichthyoplankton community (Peabody et al. 2018); the southern California
576 ichthyoplankton time series are the only data used in our analysis that pre-date the 1970s. The
577 Peabody et al. (2008) study identified several species that caused the 1960s shift and their study
578 included a much broader suite of ichthyoplankton species than our study which limits our ability
579 to evaluate whether the species driving the shifts are consistent among studies. Previous studies

580 have also documented a shift in response to the 1976/1977 PDO shift (e.g., McGowan et al.
581 2003, Peabody et al. 2018), but while our analyses did not. Our estimated biology trend is
582 consistent with the evidence of this regime shift, however, only ichthyoplankton time series are
583 available prior to the 1970s and there are gaps in the ichthyoplankton data between from the late
584 1960s through the 1970s (CalCOFI only sampled every third year during this time). The trend
585 estimate therefore has higher uncertainty during this period than elsewhere in the time
586 series, elsewhere in the time series. This likely limits any detection of a regime shift in the mid to
587 late 1970s. In a recent study, Litzow et al. (2020a) applied Bayesian DFA to 11 climate time
588 series and 48 biology time series in the Gulf of Alaska (GOA). While their study revealed
589 markedly elevated values in a shared climate trend between 2014–2019, it did not provide

590 evidence of a shift to a novel mean state in shared trends of ecosystem variability during the
591 extreme warming event. Their analysis did, however, capture the rapid community change in the
592 GOA following the 1970's PDO shift in the northeast Pacific Ocean. We note that our study, as
593 well as Litzow et al. (2020a), are based on time series that only extend a few years beyond the
594 heatwave. As additional years of data become available, the CCE and GOA Bayesian DFA
595 models could reveal different outcomes. However, this is unlikely given that the taxa and life
596 stages used in both studies are known to respond quickly to changes in ocean conditions, i.e.,
597 within less than one year, and given our assumption that the surveys are consistently sampling at
598 the right time and location to fully characterize the short term response.

599 The CCE biology time series included in this study showed strong coherence in
600 community signal in response to regional climate perturbations. Although they span across
601 multiple trophic levels, life-history strategies, and datasets, most of the biological time series
602 loaded in the same direction on the shared trend (Fig. 4b). In addition, our the CCE shared
603 biology trend and loadings captured an unusual aspect of the 2014–2016 warming events: the
604 abundance of several taxa, including young-of-year rockfish and anchovy, was high during the
605 marine heatwave (Santora et al. 2017, Thompson et al. 2019, Schroeder et al. 2019). By contrast,
606 their abundance was greatly reduced in most previous warm events, including two of the
607 strongest El Niño events on record (1982–1983, 1997–1998) and unusually low productivity
608 conditions (2005–2006, Peterson et al. 2006). High abundances of young-of-year rockfish and
609 groundfish, squid, and krill in the CCE are generally associated with more southward transport
610 and subarctic source waters, while abundances are typically far lower in years with more
611 subtropical waters, which are often associated with El Niño and anomalous warm events
612 (Ralston et al. 2015, Schroeder et al. 2019). The unexpectedly high abundance of these taxa in

613 2014–2016, despite surface-oriented marine heatwave, may be related to the prevalence of
614 subsurface waters that were more subarctic than subtropical in origin (Schroeder et al. 2019) and
615 to some concurrent strong upwelling events, particularly in spring 2015 (Peterson et al. 2015,
616 Ryan et al. 2017).

617 Our results were consistent with recent studies of several top predators in the CCE. The
618 DFA trends and loadings indicate a negative response of sea lion pup growth and weight to the
619 2014–2016 marine heatwave, which also aligns with past work showing that sea lion pup
620 condition covaries with abundance of forage such as larval anchovy and sardine, which provide
621 quality prey to sustain lactation in nursing mothers (McClatchie et al. 2016). Pup condition also
622 improved at the tail end of the marine heatwave when, despite the warm water, anchovy
623 abundance increased dramatically (Thompson et al. 2019). The trends and loadings suggest that
624 the reproductive success of some seabirds in the central CCE was not diminished by the
625 heatwave, although the heatwave had severe impacts seabird productivity in regions to the north
626 (Piatt et al. 2020).

627
628 *Regional comparison of the marine heatwave's effect on community state*
629 A compelling outcome of our analysis and a—the similar analysis of applied to Alaskan
630 species by Litzow et al. (2020a) is that neither detected state changes in North Pacific
631 communities following the massive 2014—2016 marine heatwave, despite the extremely
632 anomalous physical conditions throughout most of the basin and a litany of concurrent
633 biological, ecological, social and economic effects (see Introduction). An important
634 characteristic of both studies is the temporal scale of community analysis (1972—2017 for the
635 Gulf of Alaska (GOA) OA and 1951—2017 for the CCE). This long temporal scale provides an

Formatted: Indent: First line: 0", Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Font: (Default) Times, Italic, Font color: Black

636 important context for comparing contemporary change with the magnitude of historical
637 community shifts. In addition, the Bayesian DFA accounts for uncertainty in the shared trends in
638 a way that prevents premature detection of wholesale ecosystem shifts.

639 We note that Suryan et al. (2021), fitted a single-trend, non-Bayesian DFA model to a
640 larger set of GOA biological time series ($n = 187$) over a shorter time span (2010–2018) and
641 found evidence of a well-resolved shift that implied different community states during 2010–
642 2014 and 2015–2018. The different conclusions of Suryan et al. (2021) and Litzow et al.
643 (2020a) studies speak to an inherent tension in retrospective analyses of community change.

644 Limited time series availability means that analyses can be taxonomically and functionally broad
645 (e.g., Suryan et al. 2021), or temporally extensive (e.g., Litzow et al. 2020a), but not both. Each
646 approach has advantages, but direct comparison between the two is difficult. Given the impacts
647 of the 2014–2016 event, and that long-term warming combined with marine heatwaves will
648 push the CCE into novel climate states, we must consider ecological mechanisms that might
649 explain why these communities were apparently resilient to the marine heatwave, along with
650 revisiting methodological details that could further clarify our results.

651

652 *Environmental covariates*

653 The CCE shared biology trend and loadings captured unexpected patterns in the
654 community response to regional climate perturbations. The biology time series included in this
655 study showed strong coherence in community signal in response to regional climate
656 perturbations across multiple trophic levels, life history strategies, and datasets. Most of the
657 biological time series loaded in the same direction on the shared trend, with only a few time
658 series showing the opposite pattern. In addition, several species that are typically associated with

Formatted: Indent: First line: 0"

Formatted: Font: Italic, Font color: Black

659 cooler ocean conditions show increased relative abundances during the marine heatwave. For
660 example, past studies have shown that high abundances of pelagic juvenile rockfish and
661 groundfish, squid, and krill are associated with more southward transport in the California
662 Current, and more subarctic source waters, with abundance far lower in years with more
663 subtropical waters, which are often associated with El Niño and anomalous warm events
664 (Ralston et al. 2015, Schroeder et al. 2019). Our analysis captured this documented pattern,
665 showing a reduced abundance of these species that have been associated with subarctic source
666 waters and strong equatorward flow in the California Current during two of the strongest El Niño
667 events on record (1982–1983, 1997–1998) and unusually low productivity conditions (2005–
668 2006, Peterson et al. 2006). Our analysis also captured the unexpected increase in abundance of
669 these taxa in 2014–2016 despite the anomalously warm ocean conditions at that time. This may
670 be partially explained by the observation that subsurface waters were more subarctic, rather than
671 subtropical in origin (Schroeder et al. 2019) and that there was some strong upwelling during the
672 marine heatwave, particularly in spring 2015 (Peterson et al. 2015, Ryan et al. 2017). The
673 reduced production of juvenile/adult Pacific sardine and increase in juvenile northern anchovy
674 between 2014 and 2016 shown here and elsewhere (Thompson et al. 2019) were consistent with
675 a history of observations indicating that these species tend to respond asynchronously to ocean
676 conditions (MacCall 1996, Schwartzlose et al. 1999, Chavez et al. 2003, Deyle et al. 2013,
677 Sydeman et al. 2020). However, the responses were of opposite sign from past observations of
678 increases in sardines and decreases in anchovy under warm conditions (Chavez et al. 2003). The
679 DFA trends and loadings indicate a negative response of sea lion pup growth and weight to the
680 2014–2016 marine heatwave, which also aligns with past work showing that reduced prey
681 availability for nursing mother sea lions is unfavorable for sea lion pups (McClatchie et al.

682 2016). Notably, sea lion pup condition covaried with abundance of anchovy and sardine, which
683 provide quality prey to sustain lactation. Pup condition also improved at the tail end of the
684 marine heatwave when, despite the warm water, anchovy abundance increased dramatically and
685 by 2017 all metrics of pup condition were above average (Thompson et al. 2019). In addition,
686 the trends and loadings suggest that the reproductive success of some seabirds in the central
687 region of the CCE was not diminished by the heatwave, although recent studies have
688 documented the severe impact of the heatwave on seabird productivity in regions to the north
689 (Piatt et al. 2020). Our analysis is not able to capture the irruptions of fishes and invertebrates
690 that are rare in long term CCE time series. Many sporadically occurring taxa such as pelagic red
691 crabs (*Pleuroncodes planipes*), which are absent from California waters in most years but
692 abundant infrequently in warm years with anomalous transport from the south, are not well
693 suited to include as time series due to large number of zero observations in the survey data.

694 —————

695 Incorporating climate information in DFA models of CCE biology indicated that included that
696 models with a climate covariate performed better than models without one. The model results
697 suggest that Nitrate flux into the surface mixed layer (BEUTI) was the best-performing
698 covariate for individual species in addition to the shared trend in the southern and central CCE
699 over the past three decades. Nitrate flux had a strong positive effect on the abundance of krill and
700 some larval fishes and on the reproductive success of seabirds, and a moderate positive effect on
701 some pelagic juvenile fishes, squid, and sea lion pup births. Stronger upwelling magnitude
702 (CUTI), which is correlated with nitrate flux, was the second-best predictor of community
703 variability and had a positive albeit weaker effect on the same suite of species (S4 Fig.). These
704 findings are consistent with mechanistic understanding as upwelling increases the supply of

Formatted: Don't adjust space between Latin and Asian text,
Don't adjust space between Asian text and numbers

705 nutrients from deep to shallow waters and enhances the productivity at the lower trophic levels,
706 including juvenile rockfishes (Ralston et al. 2013), which affects foraging conditions for higher
707 trophic level species, such as seabirds (e.g., Wells et al. 2008).
708 _____ that affect foraging conditions for higher trophic level species in coastal habitats. For
709 example, past work has shown that rockfish abundance is positively associated with upwelling
710 strength (Ralston et al. 2013), and seabird productivity has been shown to increase when ocean
711 conditions support high primary production and higher trophic level prey (e.g. Wells et al. 2008).
712 Rockfish production is also higher when subsurface source waters in the California Current are
713 more subarctic rather than subtropical, which may explain the negative effect of isothermal layer
714 depth (ILD) and temperature on some rockfish species (S5 Fig., S6 Fig.). More subtropical
715 source waters and surface warming tend to be associated with deeper isopycnals, higher
716 stratification, and shallower ILD (Schroeder et al. 2019, Bograd et al. 2019). Nitrate flux
717 (BEUTI and CUTI) had a strong, negative correlations with juvenile/adult Pacific sardine and
718 larval northern anchovy and upwelling strength (CUTI) did as well. The relative abundance of
719 Pacific sardine in coastal waters off of Central California has been shown to be lower during
720 periods of strong upwelling (Santora et al. 2014, Ralston et al. 2015). This trend may reflect a
721 change in the production of Pacific sardine or a shift in their relative distribution. In addition, a
722 negative relationship between upwelling and nitrate flux and sardine recruitment can generally
723 be explained by the transfer of fish larvae to offshore areas where they have low chance of
724 survival during periods of strong equatorward flow and upwelling (Bailey and Francis 1985,
725 Nieto et al. 2014), although our understanding of the mechanisms driving anchovy population
726 dynamics is limited (Sydeman et al. 2020).

727 We acknowledge that Climate drivers often act in concert to influence community
728 variability, and here we are evaluating the effects of the climate variables one at a time. An
729 important next step of this work will be to examine whether including multiple climate
730 covariates in the CCE biology model further improves the forecast skill of the model-CCE
731 biology and our community state indicator. However, the individual climate variables are
732 collinear and share information, which affects our ability to makes inference on the covariates.
733 Furthermore, our study is an attempt to synthesize a broader suite broad synthesis of community
734 indicators and their response to climate perturbations, and therefore should not be interpreted as
735 replacing or simplifying more detailed investigations into the drivers and mechanistic
736 understanding of the indicators included here.

737

738 *Community state forecasting skill*

739 Our approach for creating simultaneous predictions of species responses and shared
740 ecosystem variability to ocean conditions shows promise for developing near-term forecasts of
741 community state. Our forecasts are based on ocean model data derived from the CCE
742 ROMS, which have been applied in several recent studies used to examine how oceanographic
743 processes affect fish recruitment variability (Tolimieri et al. 2018, Haltuch et al. 2020) and
744 productivity (Siegelman-Charbit et al. 2016), species habitat suitability (Abrahms et al. 2018,
745 Cimino et al. 2020), and species spatial distributions (Muhling et al. 2019, 2020). The CCE
746 ROMS also supports nowcasts of species distributions based on observed ocean conditions that
747 can help resource managers and users manage risks associated with fisheries bycatch and ship-
748 strike (Hazen et al. 2017, 2018, Welch et al. 2019). Moreover, multiple efforts are underway in
749 the CCE and other coastal systems to use ROMS outputs to develop short-term forecasts (1–24

Formatted: Normal, Indent: First line: 0", Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Formatted: Font: (Default) Times

750 months) of ocean conditions for uptake by scientists, managers, and other end-users (Siedlecki et
751 al. 2016, Kaplan et al. 2016, Jacox et al. 2020, Malick et al. 2020). Here, we were able to create
752 forecasts of community state and several individual species parameters one year in advance
753 based on observations of a single climate variable (nitrate flux). Forecast lead times could ~~The~~
754 ~~lead time of such a forecast could~~ be extended further by using forecasts of ocean conditions
755 rather than observed conditions, and ocean temperatures in the CCE can be skillfully forecast
756 months to a year in advance, with particularly high skill in the late winter and spring (Jacox et al.
757 2019). ~~As mentioned above, F~~ future extensions of our work will evaluate whether different a
758 combinations of climate variables ~~as well as~~ and time lags might improve our forecasting skill.

759 Using DFA to forecast attributes of community structure in the CCE allows us to create
760 simultaneous forecasts of trends, or ‘ecosystem state’, and raw time series. This~~Our~~ approach
761 could also be applied individually to each dataset in our analysis to generate taxa-specific
762 indicators (e.g., seabird productivity, juvenile fish abundance), though ~~these~~ forecasts would be
763 expected to differ from those with the entire CCE dataset. Similarly, if ecosystem states were not
764 a focus of inference, alternative forecast models could be applied (e.g., ARIMA or non-
765 parametric models, Ward et al. 2014). Forecasts for individual time series from the DFA models
766 used here can be seen as a mixture of the AR forecast on the estimated trends (Fig. 6), and linear
767 effects of forecasted climate variables on each time series (Fig. 5). Species that have strong
768 associations or loadings on the trend and estimated climate effects that are large in magnitude
769 (e.g., market squid, Pacific sanddabs, shortbelly rockfish *Sebastodes jordani*) are expected to have
770 the most accurate predictions, while those species with weak loadings and weaker effects of
771 climate variables (e.g., California smoothtongue (*Leuroglossus stictus*) are expected to have
772 worse poorer forecast performance.

Formatted: Adjust space between Latin and Asian text,
Adjust space between Asian text and numbers

773
774 In the CCE, nonstationary Nonstationary relationships are an important consideration for
775 producing reliable ecological forecasts in this ecosystem. While the year-to-year variability in
776 the estimated trend did appear to be stationary for in our community models (Fig. 4, 6), the
777 autocorrelation appeared to be nonstationary with the lag-1 autocorrelation between 2000–
778 present being significantly higher (0.82) than over the years 1981–2000 (0.23). In addition to
779 nonstationary variance parameters, future analyses may also consider nonstationary relationships
780 in the covariate relationships, or potential interactions between covariates. A growing number of
781 retrospective analyses have revealed nonstationary relationships among climate and individual
782 species or community-level variables (Puerta et al. 2019, Litzow et al. 2018, 2019, 2020 a,b,c).
783 In the northeast Pacific Ocean, these studies have been mostly focused on Alaskan ecosystems,
784 which have with long time series series that support robust statistical analysis of nonstationarity
785 in describing climate and biological systemsprocesses. The best-documented instance of
786 nonstationary relationships among climate and biology time series in the North Pacific centers on
787 a climate shift in the late 1980s (Litzow et al. 2020b). Decades of observational data on either
788 side of that event allow for statistically robust tests for nonstationarity that are not yet available
789 for post-2014–2016 conditions. The problem of having few years of available data is an
790 inescapable limitation for evaluating the possibility that nonstationary relationships might
791 accompany emerging, novel climate conditions. Similarly, the considerable historical and
792 paleoceanographic evidence for low frequency variability in the productivity of many key marine
793 populations, particularly coastal pelagic species such as Pacific sardine and northern anchovy,
794 also complicates the challenges associated with detecting nonstationarity with data of a limited
795 temporal duration (Baumgartner et al. 1992, MacCall 1996, Field et al. 2009), although such

796 patterns may lead to greater risk of “detecting” nonstationarity when in fact it may not exist
797 (Cohn and Lins 2005, Milly et al. 2015). However, early indications from Alaska suggest the
798 possibility that long-standing relationships between leading climate modes and individual
799 climate and biology time series may have changed following 2014 (Litzow et al. 2020c).

800

801 Management application

802 Using DFA to forecast attributes of community structure in the CCE allows us to create
803 simultaneous forecasts of trends, or ‘ecosystem state’, and raw time series. This approach could
804 also be applied individually to each dataset in our analysis to generate taxa-specific indicators
805 (e.g., seabird productivity, juvenile fish abundance), though forecasts would be expected to differ
806 from those with the entire CCE data. Similarly, if ecosystem states were not a focus of inference,
807 alternative forecast models could be applied (e.g., ARIMA or non parametric models, Ward et al.
808 2014). Forecasts for individual time series from the DFA models used here can be seen as a
809 mixture of the AR forecast on the estimated trends (Fig. 6), and linear effects of forecasted
810 climate variables on each time series (Fig. 5). Species that have strong associations or loadings
811 on the trend and estimated climate effects that are large in magnitude (e.g., market squid, Pacific
812 sanddabs, shortbelly rockfish *Sebastodes jordani*) are expected to have the most accurate
813 predictions, while those species with weak loadings and weaker effects of climate variables (e.g.,
814 California smoothtongue (*Leuroglossus stibius*) are expected to have worse forecast
815 performance.

816 Our approach for developing a community state indicator to track and predict the
817 response of marine ecosystems to climate perturbations has the potential to support ecosystem-
818 based and climate-ready management in multiple ways. Garnering knowledge of community

Formatted: Indent: First line: 0", Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Font: (Default) Times, Italic, Font color: Black

819 state and the potential for large shifts in ecosystem structure in response to intense and novel
820 climate perturbations can help inform better, more rapid management decisions for mitigating
821 ecological and socioeconomic impacts. Our intention is to continually update our analyses when
822 new data become available to provide the most up-to-date information on the CCE community
823 state for scientists, managers, and stakeholders.

824 The combination of long-term monitoring surveys and data with the modeling framework
825 we advance here can also help scientists identify or refine key variables of ecosystem change that
826 are summarized for ecosystem assessments in support of decision-making (Harvey et al. 2020).
827 For example, it might be prudent to emphasize ecological time series that load strongest on
828 ecosystem state trends and demonstrate strong, predictable relationships with climate variables

829 (or other covariates of interest) over time series with weaker loadings ~~on shared trends of~~
830 ~~community variability or have low or lower forecast skill with environmental variables.~~

831 Furthermore, our approach can provide valuable ecosystem information for scientific,
832 management and coastal communities during times when researchers cannot sample the biology
833 in marine ecosystems. This added value became acutely apparent in 2020 when myriad ocean
834 surveys were cancelled or limited in spatiotemporal scope due to safety restrictions associated
835 with the COVID-19 pandemic.

836 Finally, our approach provides a quantitative way to help managers discern short-term
837 periods of unusual community dynamics and/or high variability—such as the 2014–2016 marine
838 heatwave—from state shifts that represent more enduring transitions into new regimes of
839 ecosystem structure or productivity. Given that global climate change is expected to amplify
840 ocean change, approaches like the one applied here will become increasingly valuable for

841 identifying novel community states that require new marine resource management and
842 conservation considerations.

843

844 Acknowledgements

845 We thank those who have spent countless hours planning field surveys and collecting data for the
846 invaluable time series used in our study. We thank the U.S. Fish and Wildlife Service for
847 granting permission and providing resources to conduct research on the Farallon Islands National
848 Wildlife Refuge. Funders for Point Blue's Farallon Research Program include the Bently
849 Foundation, Elinor Patterson Baker Trust, Bernice Barbour Foundation, Frank A. Campini
850 Foundation, Grand Foundation, Kimball Foundation, Marisla Foundation, Giles W. and Elise G.
851 Mead Foundation, Moore Family Foundation, RHE Charitable Foundation, Volgenau
852 Foundation, and numerous individual donors. We thank B. Feist for creating the map of the
853 sampling areas (Figure 1). We thank N. Tolimieri and T.L. Rogers for their helpful comments
854 that improved this on an earlier version of this manuscript. We thank the Journal's academic
855 editor, Wei Yu, and two reviewers, Shuyang Ma and Chongliang Zhang, for their constructive
856 comments that improved our manuscript. Funding for this project came from NOAA's Fisheries
857 and the Environment (FATE) program (project 16-01) awarded to M.E.H., E.J.W., M.A.L. and
858 C.J.H. and NOAA's California Current Integrated Ecosystem Assessment program (C.J.H.). The
859 funders had no role in study design, data collection and analysis, decision to publish, or
860 preparation of the manuscript.
861 Funding for this project came from NOAA's Fisheries and the Environment (FATE) program
862 (project 16-01) and NOAA's California Current Integrated Ecosystem Assessment program.

863

Formatted: Font: (Default) Times New Roman

Formatted: Line spacing: Double

Formatted: Font: (Default) Times New Roman

864 **References**

- 865 Abrahms B, Welch H, Brodie S, Jacox MG, Becker EA, Bograd SJ et al. Dynamic ensemble
866 models to predict distributions and anthropogenic risk exposure for highly mobile species.
867 Divers. Distrib. 2019; 25: 1182-1193. doi: 10.1111/ddi.12940.
- 868 Anderson SC, Ward EJ. Black swans in space: modelling spatiotemporal processes with
869 extremes. Ecology 2019; 100: e02403. doi:10.1002/ecy.2403.
- 870 Anderson SC, Branch TA, Cooper AB, Dulvy NK. Black-swan events in animal populations.
871 Proc. Natl. Acad. Sci. U.S.A. 2017; 114: 3252–3257.
- 872 Anderson PJ, Piatt JF. [1999](#). Community reorganization in the Gulf of Alaska following ocean
873 climate regime shift. Mar. Ecol. Prog. Ser. 1999; 189: 117-123.
- 874 [Amante C, Eakins BW.ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources
875 and Analysis. NESDIS NGDC-24, National Geophysical Data Center, United States
876 Department of Commerce, Boulder, CO, March 2009. 19 p.](#) DOI: 10.7289/V5C8276M
- 877 Bailey KM, Francis RC. Recruitment of Pacific Whiting, *Merluccius productus*, and the ocean
878 environment. Mar. Fish. Rev. 1985; 47: 8-11.
- 879 Baumgartner TR, Soutar A, Ferreira-Bartrina V. Reconstructions of the history of Pacific sardine
880 and northern anchovy populations over the past two millennia from sediments of the Santa
881 Barbara Basin, California. CalCOFI Rep. 1992; 33: 24-40.
- 882 Beaugrand G, Edwards M, Brander K, Luczak C, Ibanez F. Causes and projections of abrupt
883 climate-driven ecosystem shifts in the North Atlantic. Ecol. Lett. 2008; 11: 1157–1168.
- 884 Beaugrand G, Conversi A, Chiba S, Edwards M, Fonda-Umani S, Greene C et al. Synchronous
885 marine pelagic regime shifts int he Northern Hemisphere. Phil. Trans. R. Soc. 2016; 370:
886 20130272. doi: 10.1098/rstb.2013.0272.

Formatted: Font: (Default) Times New Roman

887 Benson AJ, Trites AW. Ecological effects of regime shifts in the Bering Sea and eastern North
888 Pacific Ocean. *Fish Fish.* 2002; 3: 95–113.

889 Bjorkstedt EP, Goericke R, McClatchie S, Weber E, Watson W, Lo N et al. State of the
890 California current 2009-2010: Regional variation persists through transition from the la
891 Niña to el Niño (and back?). *CalCOFI Rep.* 2010; 51.

892 Bograd S, Schroeder ID, Jacox MG. [2019](#). A water mass history of the Southern California
893 current system. *Geophys. Res. Lett.* 2019; 46: 6690-6698.

894 Bond NA, Cronin MF, Freeland H, Mantua N. Causes and impacts of the 2014 warm anomaly in
895 the NE Pacific. *Geophys. Res. Lett.* 2015; 42, 3414–3420. doi: 10.1002/2015GL063306.

896 Bürkner P-C, Gabry J, Vehtari A. [2020](#). Approximate leave-future-out cross-validation for
897 Bayesian time series models. *J. Stat Comput. Simul.* 2020; 90: 2499-2523.

898 Brodeur RD, Auth TD, Phillips AJ. Major shifts in pelagic micronekton and zooplankton
899 community structure in an upwelling ecosystem related to an unprecedented marine
900 heatwave. *Front. Mar. Sci.* 2019. doi: 10.3389/fmars.2019.00212.

901 Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M et al. Stan: A
902 Probabilistic Programming Language. *J. Stat. Softw.* 2017. doi: 0.18637/jss.v076.i01.

903 Cavole LM, Demko AM, Diner RE, Giddings A, Koester I, Pagniello CM et al. Biological
904 impacts of the 2013–2015 warm-water anomaly in the Northeast Pacific: winners, losers,
905 and the future. *Oceanography* 2016; 29, 273–285.

906 Chavez FP, Ryan J, Lluch SE, Niquen MC. From Anchovies to Sardines and Back: Multidecadal
907 change in the Pacific Ocean. *Science* 2003; 299: 217-221.

908 [Checkley DM Jr, Barth JA. Patterns and processes in the California Current System. Prog.](#)
909 [Ocean.](#) 2009; 83, 49-64.

910 Cimino MA, Santora JA, Schroeder I, Sydeman W, Jacox MG, Hazen EL, Bograd SJ. Essential
911 krill species habitat resolved by seasonal upwelling and ocean circulation models within the
912 large marine ecosystem of the California Current System. *Ecography* 2003; 43, 1-15. doi:
913 10.1111/ecog.05204.

914 Cohn TA, Lins, HF. Nature's style: Naturally trendy. *Geophys. Res. Lett.* 2005. doi:
915 10.1029/2005GL024476

916 Deyle ER, Fogarty M, Hsieh C-h, Kaufman L, MacCall AD, Munch SB et al. Predicting climate
917 effects on Pacific sardine. *Proc Natl Acad Sci USA* 2013; 110: 6430-6435.

918 Field DB, Baumgartner TR, Ferreira V, Gutierrez D, Lozano-Montes H, Salvatteci R, Soutar A.
919 Variability from scales in marine sediments and other historical records. In: Checkley DM,
920 Alheit J, Oozeki Y, editors. *Climate change and small pelagic fish*. Cambridge: Cambridge
921 University Press; 2009. pp. 45-63.

922 Field JC, Miller RA, Santora JA, Tolimieri N, Haltuch MA, Brodeur RD et al. Spatiotemporal
923 patterns of variability in the abundance and distribution of winter-spawned pelagic juvenile
924 rockfish in the California Current. *PloS one* 2021; 16: e0251638.

925 Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian Data Analysis*. 3rd
926 ed. CRC Press; 2013.

927 Gelman A, Rubin DB. Inference from Iterative Simulation Using Multiple Sequences. *Statist.*
928 *Sci.* 1992; 7: 457-472.

929 Gottsch AD. Zoogeography of the San Andreas Fault system: Great Pacific Fracture Zones
930 correspond with spatially concordant phylogeographic boundaries in western North
931 America. Biological Reviews 2016; 91: 235-254.

932 Haltuch MA, Tolimieri N, Lee Q, Jacox MG. Oceanographic drivers of petrale sole recruitment

Formatted: Font: Times New Roman, 12 pt

Formatted: Font: Times New Roman, 12 pt

Formatted: Font: Times New Roman, 12 pt

933 in the California Current Ecosystem, Fish. Ocean. 2020; 29:122-136. doi:
934 10.1111/fog.12459.

935 Hare SR, Mantua NJ. Empirical evidence for North Pacific regime shifts in 1977 and 1989. Prog.
936 Oceanogr. 2000; 47: 103–145. doi: 10.1016/s0079-6611(00)00033-1.

937 Harvey CJ, Fisher J, Samhouri, JF, Williams GD, Francis TB, Jacobson KC et al. The
938 importance of long-term ecological time series for integrated ecosystem assessment and
939 ecosystem-based management. Prog. Oceanogr. 2020. doi: 10.1016/j.pocean.2020.102418.

940 Hazen EL, Palacios DM, Forney KA, Howell EA, Becker E, Hoover AL et al. WhaleWatch: a
941 dynamic management tool for predicting blue whale density in the California Current. J.
942 Appl. Ecol. 2017. doi: 10.1111/1365-2664.12820.

943 Hazen EL, Scales KL, Maxwell SM, Briscoe D, Welch, H, Bograd et al. A dynamic ocean
944 management tool to reduce bycatch and support sustainable fisheries. Sci. Adv. 2018, 4:
945 eaar3001.

946 Hobday AJ, Oliver ECJ, Sen Gupta A, BenthuySEN JA, Burrows MT, Donat MG et al.
947 Categorizing and naming marine heatwaves. Oceanography 2018; 31: 162 - 173.

948 Hobday AJ, Spillman CM, Paige Eveson J, Hartog JR. Seasonal forecasting for decision support
949 in marine fisheries and aquaculture. Fish. Oceanogr. 2016; 25: 45–56.

950 Hoffman MD, Gelman A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in
951 Hamiltonian Monte Carlo. J. Mach. Learn. Res. 2014; 15: 1593–1623.

952 Holmes EE. Beyond theory to applications and evaluation: Diffusions approximations for
953 population viability analysis. Ecol. Appl. 2004; 14: 1272-1293.

954 Holmes EE, Ward EJ, Scheuerell, MD. Analysis of multivariate time series using the MARSS
955 package; 2018.

956 Jacox MG, Alexander MA, Siedlecki S, Chen K, Kwon Y-O, Brodie S et al. Seasonal-to-
957 interannual prediction of North American coastal marine ecosystems: Forecast methods,
958 mechanisms of predictability, and priority developments. *Prog. Oceanogr.* 2020. doi:
959 10.1016/j.pocean.2020.102307.

960 Jacox MG, Alexander MA, Stock CA, Hervieux G. On the skill of seasonal sea surface
961 temperature forecasts in the California Current System and its connection to ENSO
962 variability. *Clim. Dyn.* 2019; 53: 7519-7533.

963 Jacox MG, Alexander MA, Mantua NJ, Scott JD, Hervieux G, Webb RS et al. Forcing of
964 multiyear extreme ocean temperatures that impacted California Current living marine
965 resources in 2016. *Bull. Am. Meteorol. Soc.* 2018a; 99: S27–S33. doi: 10.1175/BAMS-D-
966 17-0119.1

967 Jacox MG, Edwards CA, Hazen EL, Bograd, SJ. Coastal upwelling revisited: Ekman, Bakun,
968 and improved upwelling indices for the U.S. west coast. *J. Geophys. Res.* 2018b.
969 doi:10.1029/2018JC014187.

970 Jacox MG, Fletcher J, Moore AM, Edwards CA. ENSO and the California Current coastal
971 upwelling response. J. Geophys. Res. 2015. doi:10.1002/2014JC010650

972 Jones T, Parish JK, Peterson WT, Bjorkstedt EP, Bond NA, Balance LT et al. Massive mortality
973 of a planktivorous seabird in response to a marine heatwave. *Geophys. Res. Lett.* 2018; 45:
974 3193-3202.

975 Kaplan IC, Williams GD, Bond NA, Hermann AJ, Siedlecki S. et al. Cloudy with a chance of
976 sardines: forecasting sardine distributions using regional climate models. *Fish. Oceanogr.*
977 2016; 25: 15-2.

978 Koslow J, Goericke R, Watson W. Fish assemblages in the Southern California Current:

979 relationships with climate, 1951–2008. Fish. Oceanogr. 2013; 22: 207–219.

980 Koslow JA, Hobday AJ, Boehlert GW et al. Climate variability and marine survival of coho

981 salmon (*Oncorhynchus kisutch*) in the Oregon production area. Fish. Oceanogr. 2002; 11: 65-

982 77.

983 Laufkötter C, Zscheischler J, Frölicher TL. High-impact marine heatwaves attributable to human-

984 induced global warming. Science 2020; 369: 1621-1625.

985 Litzow MA, Ciannelli L. 2007. Oscillating trophic control induces community reorganization in

986 a marine ecosystem. Ecol. Lett. 2007; 10: 1124–1134. doi: 10.1111/j.1461-

987 0248.2007.01111.x.

988 Litzow MA, Ciannelli L, Puerta P, Wettstein JJ, Rykaczewski RR, Opiekun M. Non-stationary

989 climate–salmon relationships in the Gulf of Alaska. Proc. R. Soc. B Biol. Sci. 2018; 285:

990 20181855. doi: 10.1098/rspb.2018.1855.

991 Litzow MA, Ciannelli L, Puerta P, Wettstein JJ, Rykaczewski RR, Opiekun M. Nonstationary

992 environmental and community relationships in the North Pacific Ocean. Ecology 2019; 100:

993 ecy.2760. doi: 10.1002/ecy.2760.

994 Litzow MA, Hunsicker ME, Ward EJ, Anderson SC, Gao J, Zador S et al. [et al. 2020a](#).

995 Evaluating ecosystem change as Gulf of Alaska temperature exceeds the limits of

996 preindustrial variability. Prog. Ocean. 2020a; 117: 7665-7671.

997 Litzow MA, Hunsicker ME, Bond NA, Burke BJ, Cunningham CJ, Gosselin JL, Norton EL,

998 Ward EJ, Zador SG. [2020b](#). The changing physical and ecological meanings of North Pacific

999 Ocean climate indices. Proc. Natl. Acad. Sci. U.S.A. 2020b; 117: 7665–7671. doi:

1000 10.1073/pnas.1921266117.

1001 Litzow MA, Malick MJ, Bond NA, Cunningham CJ, Gosselin JL, Ward EJ. [2020c](#). Quantifying

1002 a novel climate through changes in PDO-climate and PDO-salmon relationships. Geophys.
1003 Res. Lett. 2020**cb**. doi: 10.1029/2020GL087972.

1004 MacCall AD. Patterns of low-frequency variability in fish populations of the California Current.
1005 CalCOFI Reports 1996; 37: 100-110.

1006 Malick M.J, Siedlecki SA, Norton EL, Kaplan IC, Haltuch MA, Hunsicker ME et al. [2020](#).
1007 Environmentally driven seasonal forecasts of Pacific hake distribution. Front. Mar. Sci.
1008 [2020](#); 7: 578490. doi: 10.3389/fmars.2020.578490.

1009 Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC. A Pacific interdecadal climate
1010 oscillation with impacts on salmon production. Bull. Am. Meteorol. Soc. 1997; 78, 1069–
1011 1079.

1012 McCabe RM, Hickey BM, Kudela RM, Lefebvre KA, Adams NG, Bill BD et al. An
1013 unprecedented coastwide toxic algal bloom linked to anomalous ocean conditions. Geophys.
1014 Res. Lett. 2016; 43: 10366–10376. doi:10.1002/2016GL070023.

1015 McClatchie S, Goericke R, Koslow JA, Schwing FB, Bograd SJ, Charter R-W et al. The State of
1016 the California Current, 2007–2008: La Niña conditions and their effects on the ecosystem.
1017 Cal-COFI Rep. 2008; 49: 39–76.

1018 McClatchie S, Goericke R, Schwing FB, Bograd SJ, Peterson WT, Emmett R et al. The state of
1019 the California Current, 2008–2009: Cold conditions drive regional difference. CalCOFI
1020 Rep. 2009; 50: 43–68.

1021 [McClatchie S, Duffy-Anderson J, Field JC, Goericke R, Griffith D, Hanisko DS, Hare JA,](#)
1022 [Lyczkowski-Shultz J, Peterson WT, Watson W, Weber ED. Long time series in US fisheries](#)
1023 [oceanography. Oceanography 2014; 27: 48-67.](#)

1024 McClatchie S, Field J, Thompson AR, Gerrodette T, Lowry M, Fiedler PC et al. Food limitation

Formatted: Font: Pattern: Clear (White)

1025 of sea lion pups and the decline of forage off central and southern California. *R. Soc.*
1026 *opensci.* 2016; 3: 150628. doi: 10.1098/rsos.150628.

1027 McGowan JA, Bograd SJ, Lynn RJ, Miller AJ. The biological response to the 1977 regime shift
1028 in the California Current. *Deep Sea Res. II* 2003; 50: 2567-2582.

1029 Milly PCD, Betancourt J, Falkenmark M, Hirsch, RM, Kundzewicz ZW, Lettenmaier DP et al.
1030 On critiques of “Stationarity is dead: Whither water management?” *Water Res. Res.* 2015;
1031 51: 7785–7789.

1032 Möllmann C, Diekmann R. Marine Ecosystem Regime Shifts Induced by Climate and
1033 Overfishing: A Review for the Northern Hemisphere. *Adv. Ecol. Res.* 2012; 47: 303-347.

1034 Morgan CA, Beckman BR, Weitkamp LA, Fresh KL. Recent ecosystem disturbance in the
1035 Northern California Current. *Fisheries* 2019; 44: 465-474. doi: 10.1002/fsh.10273.

1036 Muhling B, Brodie S, Jacox MG, Snodgrass O, Dewar H, Tommasi D, Edwards C, Xu Y, Snyder
1037 S, Childers J. Dynamic habitat use of albacore and their primary prey species in the
1038 California Current System. *CalCOFI Reports* 2019; 60: 79-93.

1039 Muhling B, Brodie S, Smith JA, Tommasi D, Gaitan CF, Hazen EL et al. Predictability of
1040 species distributions deteriorates under novel environmental conditions in the California
1041 Current System. *Front. Mar. Sci.* 2020; doi:10.3389/fmars.2020.00589.

1042 Neveu E, Moore AM, Edwards CA, Fiechter J, Drake P, Jacox MG, Nuss E. A historical analysis
1043 of the California Current using ROMS 4D-Var. Part I: System configuration and
1044 diagnostics, *Ocean Model.* 2016; 99: 133-151. doi:10.1016/j.ocemod.2015.11.012.

1045 Nielsen JM, Rogers LA, Brodeur RD, Thompson AR, Auth TD, Dreary AL et al. Responses of
1046 ichthyoplankton assemblages to the recent marine heatwave and previous climate
1047 fluctuations in several Northeast Pacific marine ecosystems. *Glob. Chang. Biol.* 2020; 27:

1048 506-520.

1049 Nieto K, McClatchie S, Weber ED, Lennert-Cody CE. Effect of mesoscale eddies and streamers
1050 on sardine spawning habitat and recruitment success off Southern and central California, J.
1051 Geophys. Res. Oceans 2014; 119: 6330–6339, doi:10.1002/014JC010251.

1052 Peabody CE, Thompson AR, Sax DF, Morse RE, Perretti CT. Decadal regime shifts in southern
1053 California's ichthyoplankton assemblage. Mar. Ecol. Prog. Ser. 2018; 607: 71-83.

1054 Peterson WT, Emmett R, Goericke R, Venrick E, Mantyla A, Bograd S et al. The State of the
1055 California Current, 2005-2006: warm in the north, cool in the south. CalCOFI Reports
1056 2006; 47: 30–74.

1057 Peterson WT, Fisher JL, Strub PT, Du X, Risien C, Peterson J et al. [2015-The pelagic ecosystem](#)
1058 [in the Northern California Current off Oregon during the 2014-2016 warm anomalies](#)
1059 [within the context of the past 20 years. J. Geophys. Res. 2017; 122: 7267-7290.](#)
1060 <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017JC012952>

1061 Piatt JF, Parrish JK, Renner HM, Schoen SK, Jones TT, Arimitsu ML et al. Extreme mortality
1062 and reproductive failure of common murres resulting from the northeast Pacific marine
1063 heatwave of 2014-2016. PLoS ONE 2020; 15: e0226087. doi:
1064 10.1371/journal.pone.0226087

1065 Planque B, Arneberg P. [2018- Principal component analyses for integrated ecosystem](#)
1066 [assessments may primarily reflect methodological artefacts. ICES J. Mar. Sci. 2018; 75:](#)
1067 [1021–1028.](#)

1068 Puerta P, Ciannelli L, Rykaczewski R, Opiekun M, Litzow MA. Do Gulf of
1069 Alaska fish and crustacean populations show synchronous non-stationary responses
1070 to climate? Prog. Oceanogr. 2019; 175: 161–170. doi: 10.1016/j.pocean.2019.

Formatted: Font: Font color: Text 1, Do not check spelling or grammar

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Indent: Left: 0", Hanging: 0.33", No widow/orphan control

1071 04.002.

1072 R Core Team. R: A language and environment for statistical computing. R

1073 Foundation for Statistical Computing, Vienna, Austria, 2021.

1074

1075 Ralston S, Sakuma KM, Field JC. Interannual variation in pelagic juvenile

1076 rockfish (*Sebastodes* spp.) abundance – going with the flow. *Fish. Ocean.* 2013; 22:288–308. doi: 10.1111/fog.12022.

1077

1078 Ralston S, Field JC, Sakuma KM. Long-term variation in a central California pelagic forage

1079 assemblage. *J. Mar. Sys.* 2015; 146: 26–37. doi: 10.1016/j.jmarsys.2014.06.013.

1080 Ryan JP, Kudela RM, Birch JM, Blum M, Bowers HA, Chavez FP et al. Causality of an extreme

1081 harmful algal bloom in Monterey Bay, California, during the 2014–2016 northeast Pacific

1082 warm anomaly. *Geophys. Res. Lett.* 2017; 44: 5571–5579. doi: 10.1002/2017GL072637.

1083

1084 Sakuma KM, Field JC, Mantua NJ, Ralston S, Marinovic BB, Carrion CN.

1085 Anomalous epipelagic micronekton assemblage patterns in the neritic waters of the California

1086 Current in spring 2015 during a period of extreme ocean conditions. *CalCOFI Reports*

1087 2016; 57: 163–183.

1088 Sanford E, Sones JL, García-Reyes M, Goddard JH, Largier JL. Widespread shifts in the coastal

1089 biota of northern California during the 2014–2016 marine heatwaves. *Sci. Rep.* 2019; 9: 1–

1090 14.

1091 Santora JA, Mantua NJ, Schroeder ID, Field JC, Hazen E, Bograd SJ et al. Habitat compression

1092 and ecosystem shifts as potential links between marine heatwave and record whale

1093 entanglements. *Nat. Commun.* 2020; 11: 1–12.

1094 Santora JA, Hazen EL, Schroeder ID, Bograd SJ, Sakuma KM, Field JC. Impacts of ocean
1095 climate variability on biodiversity of pelagic forage species in an upwelling ecosystem.
1096 Mar. Ecol. Prog. Ser. 2017; 580: 205-220.

1097

1098 Santora JA, Schroeder ID, Field JC, Wells BK, Sydeman WJ. [2014](#). Spatiotemporal
1099 dynamics of ocean conditions and forage taxa reveals regional structuring of predator-prey
1100 relationships. Ecol. Appl. 2014; 24:1730-1747. doi:10.1890/13-1605.1.

1101 Schroeder ID, Santora JA, Bograd SJ, Hazen EL, Sakuma, KM, Moore AM et al. Source water
1102 variability as a driver of rockfish recruitment in the California Current Ecosystem:
1103 implications for climate change and fisheries management. Can. J. Fish. Aquat. Sci. 2019;
1104 76: 950-960. doi: 10.1139/cjfas-2017-0480.

1105 Schwartzlose RA, Alheit J, Bakun A, Baumgartner TR, Cloete R, Crawford RJM et al. [1999](#).
1106 Worldwide large-scale fluctuations of sardine and anchovy populations. S. Afr. J. Mar. Sys.
1107 1999; 21: 289-347.

1108 Sen Gupta A, Thomsen M, BenthuySEN JA, Hobday AL, Oliver E, Alexander LV, et al. Drivers
1109 and impacts of the most extreme marine heatwaves events. Sci. Rep. 2020; 10: 19359. doi:
1110 10.1038/s41598-020-75445-3.

1111 Seo H, Brink KH, Dorman E, Koracin D, Edwards CA. What determines the spatial pattern in
1112 summer upwelling trends on the US West Coast? J. Geophys. Res. Oceans 2012. doi:
1113 10.1029/2012JC008016.

1114 Siedlecki SA, Kaplan IC, Hermann AJ, Nguyen TT, Bond NA, Newton JA et al. Experiments
1115 with seasonal forecasts of ocean conditions for the northern region of the California
1116 Current upwelling system. Sci. Rep. 2016; 6: 27203. DOI: 10.1038/srep27203.

1117 Siegelman-Charbit L, Koslow JA, Jacox MG, Hazen EL, Bograd SJ, Miller EF, McGowan JA.
1118 Physical forcing on fish abundance in the southern California Current System. Fish. Ocean.
1119 2018; 27: 475–488. doi:10.1111/fog.12267.

1120 Stan Development Team. RStan: The R interface to Stan. 2018.

1121 Suryan RM, Arimitsu ML, Coletti HA, Hopcroft RR, Lindeberg MR, Barbeaux SJ et al.
1122 Ecosystem response persists after a prolonged marine heatwave. Sci. Rep. 2021; 11: 6235.

1123

1124 Sydeman WJ, Dedman S, Garcia-Reyes M, Thompson SA, Thayer JA, Bakun A, MacCall AD.
1125 Sixty-five years of northern anchovy population studies in the southern California Current:
1126 a review and suggestion for sensible management. ICES J. Mar. Sci. 2020; 77: 486–499.

1127 Thompson AR, Hyde JR, Watson W, Chen DC, Guo LW. Rockfish assemblage structure and
1128 spawning locations in southern California identified through larval sampling. Mar. Ecol.
1129 Prog. Ser. 2016; 547: 177-192.

1130 Thompson AR, Schroeder ID, Bograd SJ, Hazen EL, Jacox MG, Leising AL et al. State of the
1131 California Current 2018-19: a novel anchovy regime and a new marine heatwave?
1132 CalCOFI Reports 2019; 60: 1-65.

1133 Tolimieri N, Haltuch M, Lee Q, Jacox MG, Bograd SJ. Oceanographic drivers of sablefish
1134 recruitment in the California Current. Fish. Ocean, 2018; 27: 458-474,
1135 doi:10.1111/fog.12266.

1136 Tommasi D, Stock CA, Hobday AJ, Methot R, Kaplan IC, Eveson JP et al. Managing living
1137 marine resources in a dynamic environment: the role of seasonal to decadal climate
1138 forecasts. Prog. Oceanogr. 2017; 152: 15–49.

← Formatted: Indent: Left: 0", Hanging: 0.38"

1139 Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-
1140 validation and WAIC. *Statistics and Computing* 2017; 27: 1413-1432. doi:
1141 10.1007/s11222-016-9696-4.

1142 Walker Jr HJ, Hastings PA, Hyde JR, Lea RN, Snodgrass OE, Bellquist LF. Unusual occurrences
1143 of fishes in the Southern California Current System during the warm water period of 2014–
1144 2018. *Estuar. Coast. Shelf Sci.* 2020; 236: 106634.

1145 Walsh JE, Thoman RL, Bhatt US, Bieniek PA, Brettschneider B, Brubaker M et al. The high
1146 latitude heat wave of 2016 and its impacts on Alaska. *Bull. Am. Meteorol. Soc.* 2018; 99:
1147 S39–S43. doi: 10.1175/BAMS-D-17-0105.

1148 Ward EJ, Holmes EE, Thorson JT, Collen B. Complexity is costly: a meta-analysis of parametric
1149 and non-parametric methods for short-term population forecasting. *Oikos* 2014; 123: 652–
1150 661.

1151 Ward EJ, Anderson SC, Damiano LA, Hunsicker ME, Litzow MA. Modeling regimes with
1152 extremes: the bayesdfa package for identifying and forecasting common trends and
1153 anomalies in multivariate time-series data. *R J* 2019; 11: 46–55.

1154 Ward EJ, Anderson SC, Damiano LA, Malick MJ. bayesdfa: Bayesian Dynamic Factor Analysis
1155 (DFA) with 'Stan'. R package version 1.1.0 2020. [https://CRAN.R-
1156 project.org/package=bayesdfa](https://CRAN.R-project.org/package=bayesdfa)

1157 Welch H, Hazen EL, Briscoe DK, Bograd SJ, Jacox MG, Eguchi T et al. Environmental
1158 indicators to reduce loggerhead turtle bycatch offshore of Southern California. *Ecol. Ind.*
1159 2019; 98: 657–664. doi:10.1016/j.ecolind.2018.11.001.

1160 Wells BK, Field JC, Thayer JA, Grimes CB, Bograd SJ, Sydeman WJ et al. Untangling the
1161 relationships among climate, prey and top predators in an ocean ecosystem. *Mar. Ecol.*

1162 Prog. Ser. 2008; 364: 15-29.
1163 Wernberg T, Bennett S, Babcock RC, De Bettignies T, Cure K, Depczynski M et al. Climate-
1164 driven regime shift of a temperate marine ecosystem. Science 2016; 353: 169-172.
1165 Zuur AF, Tuck ID, Bailey N. Dynamic factor analysis to estimate common trends in fisheries
1166 time series. Can. J. Fish. Aquat. Sci. 2003; 60: 542–552.

1167

1168
1169
1170
1171

1172
1173

1174 **Supporting Information**

1175 **S1 Appendix:** Standardization of time series from spatially resolved datasets.
1176 **S1 Table:** Climate and biology time series included in the analyses
1177 **S2 Table:** Summary information for the Bayesian DFA biology-covariate and biology only
1178 models (years 1981–2017).
1179 **S3 Table:** Observations, predictions, and prediction errors for single species parameters in 2018.
1180 **S1 Figure:** Climate and biology time series used in the study analyses.
1181 **S2 Figure:** AR(1) coefficient on the southern/central California latent climate trend and support
1182 for a heavy-tailed deviations of the latent trend.
1183 **S3 Figure:** The Student-t deviations degrees of freedom parameter (ν) in the southern/central
1184 California biology trend.

1185 **S4 Figure:** A summary of the effect of the Cumulative Upwelling Transport Index (CUTI) on
1186 the individual single species parameter included in the DFA analyses

1187 **S5 Figure:** A summary of the effect of the Isothermal Layer Depth (ILD) on the individual
1188 single species parameter included in the DFA analyses.

1189 **S6 Figure:** A summary of the effect of the sea surface temperature on the individual single
1190 species parameter included in the DFA analyses

1191 **S7 Figure:** Community variability in the southern California Current ecosystem (1981–2018).

1192 **S8 Figure:** Forecasts and model estimates of the ‘true’ community state in the southern and
1193 central California Current in years 2009–2018.

1194 **S89 Figure:** Fitted values for biology-covariate model including BEUTI (nitrate flux) as a
1195 covariate (1981–2017).

1196 **S910 Figure:** Log coefficient of variation (CV) of 2018 predictions of individual species
1197 parameters plotted against the mean and log CV of loadings related to each species, and the
1198 mean and log CV of coefficients relating each species to BEUTI (nitrate flux).

Dear Dr. Wu,

Thank you for giving us the opportunity to revise our manuscript for PLOS Climate. My co-authors and I have addressed all of the reviewers' comments, as well as the notes from the editorial team regarding the journal's requirements. You can find our responses below.

We think the reviewers' suggestions were constructive for clarifying our study approach, methodology and results, and we accepted most of them. We note, however, that we were less inclined to remove/re-organize text pertaining to the forecasting component of the paper, as suggested by reviewer #2. The forecasting component is a novel development of our modeling approach and therefore we think it's appropriate to include most of the existing text. We believe this information is valuable for readers who are interested in applying a similar approach to their own study systems.

Sincerely,
Mary Hunsicker (on behalf of all co-authors)

Reviewer #1: Climate-induced community responses are highly concerned as the urgent need of ecosystem-based and climate-ready fisheries management. Especially, marine heatwaves such extreme events have largely affected marine organisms and further the communities, and will lead to community reorganizations and the so-called regime shifts, results in great economic and social effects. This study focuses on the California Current Ecosystem (CCE), a famous upwelling ecosystem, using the newly developed Bayesian Dynamic Factor Analysis (DFA) to track the community's response to climatic drivers and to forecast future changes in the community state. There are some insightful results to benefit understanding the biological responses to the unparalleled marine heatwave in 2014-2016 in the CCE. The forecasting ability and its effectiveness of this methods are proved by this study. Hence the methods can be extended to other areas in future works. The manuscript is well-structured and well-writing. I only have some minor comments for the authors to make corresponding revisions.

Response: Thank you for reviewing our manuscript and providing us with constructive feedback.

*Please note: All line numbers in our responses below refer to the revised manuscript with tracked changes.

Minor comments

1. Lines 62. A simple definition on marine heatwave is recommended.

Response: We added a definition to the introductory paragraph of the manuscript (lines 63-64).

These effects may be exacerbated when changes in ocean conditions are more extreme, such as during marine heatwaves (prolonged events of anomalously warm ocean waters).

2. Lines 74-77. It is better to clarify that the thermal shifts that from a cold to a warm regime happened in the northeast Pacific.

Response: We added this information to the text (line 76)

3. Lines 79-85. Providing the exact numbers of SST or heat content anomalies would be helpful for readers to understanding the unparalleled marine heatwave.

Response: We added this information to the text (line 82-84).

However, between 2014 and 2016 these ecosystems experienced a marine heatwave that involved the warmest sea surface temperature (SST) and heat content anomalies that had ever been observed over large areas of the North Pacific, with SST anomalies over 6°C (Bond et al. 2015; Walsh et al. 2018).

4. Lines 120-129. The authors describe the DFA and the Bayesian DFA that both are useful replacements of the traditional PCA. However, the superiority of the Bayesian DFA to the DFA should be further clarified.

Response: Thanks, we edited this sentence to indicate that the Bayesian approach offers added flexibility (lines 135-140).

Ward et al. (2019) recently developed a Bayesian implementation of DFA that offers added flexibility in model aspects over conventional approaches; examples include allowing for extreme “black swan” events (rare and difficult to predict events; Anderson et al. 2017), and trend processes that don’t follow a random walk. Output from these Bayesian DFA models can also be used to estimate the probability of extreme events occurring or switches among contrasting system states.

5. Lines 151-152. I wonder why the authors use the same climate variables from two regions. Are they distinguished from each other in the long-term scale?

Response: We used the climate variables from the two regions because the regions differ in terms of their biogeography. There is a well-documented biogeographical boundary at Point Conception, California and we wanted to account for this in our analysis. We added this information to the manuscript on lines 190-193.

In the alongshore direction, we calculated averages for two regions with a division at Point Conception, California, separating the southern portion of the CCE ($31\text{--}34.5^{\circ}\text{N}$) from the central region ($34.5\text{--}40.5^{\circ}\text{N}$, Fig. 1). *This is in response to the recognition of Point Conception as a major biogeographic boundary for the California Current System, with differing wind and current patterns north and south of that feature (Checkley and Barth 2009, Gottscho 2016).*

Our results indicate that the time series from both regions are in agreement with respect to their loading on a single climate trend.

6. Lines 166-169 There should be literature to support that the ENSO is the dominant mode of interannual variability influencing the CCE. In addition, why do not the authors consider the effects of decadal-scale climatic variability modes such as the PDO (as mentioned in the introduction that the PDO is important for the northeast Pacific) and use winter averages?

Response: We now cite Jacox et al. 2015 to support our statement about ENSO being the dominant mode of interannual variability influencing the CCE. The decadal variability (e.g., PDO) is captured in the variables we use, but we did not include the PDO index as a predictor in our analysis because we wanted to focus on the proximate drivers of ecosystem change. Additionally, using ROMS outputs will enable us to use ROMS forecasts to then forecast biological changes in the CCE.

Jacox MG, Fletcher J, Moore AM, Edwards CA. ENSO and the California Current coastal upwelling response. *J. Geophys. Res. Oceans* 2015.

7. Lines 190-191. How do the authors deal with the unequal time span in biological variables? Will it cause or not cause problems in the Bayesian DFA?

Response: DFA can handle unequal time spans so this is not a problem for our biology model. Although unequal time span may increase uncertainty in the latent trend during years with less overlap in time series because there is less data. The uncertainty associated with the greater paucity of data in the early part of the biology time series included in our study is more explicitly discussed on lines 471-474 and 638-648.

8. Lines 249-250. Similar to question 5, as biological variables are pooled together, is there any need to use climate variables with spatial differences (e.g. SST in the central and southern regions)?

Response: The biological variables are pooled together in the biology DFA models; however, because they are sampled across a biogeographical boundary (Point Conception, California; see

response to #5) we felt that we should include climate time series from both regions (south and central).

9. Lines 253-257. Are they the same between LOO-CV and k-fold cross validation in which the k is equal to the number of years?

Response: k is equal to the total number of years for the LOO-CV, which was used for regime detections in DFA trends. However, k is equal to 10 years (last 10 years of the dataset) for the LFO-CV, which was used to identify the best climate, biology, and biology-climate DFA models.

10. Lines 253-277. The LOO-CV, k-fold cross validation and LFO-CV confuse me. To my knowledge, the authors use LFO-CV to firstly identify parameters of (1), (2), (3) and (4), secondly identify the most appropriate error structure for the climate dataset. If it is, the authors should describe more on the LFO-CV rather than introduce other types of information criteria in this part.

Response: Yes, this is mostly correct. We only used LOO-CV to identify regimes/alternating community states in the DFA trends (3). Here we describe the LOO-CV because it is a commonly used method and we want to distinguish this method from the LFO-CV method. We also use the LOO-CV to identify regimes in the DFA trends later in the methods. We cite the Burkner et al. paper for more details on the LFO-CV method.

11. Lines 291-293. It should be the number of the regimes instead of trends.

Response: Thanks for catching this. We corrected this typo.

12. Lines 306. The biology-climate model has only one climate covariate which I think may be incomplete. In my opinion, it would be better to include all the climate variable into the model and then conduct a variable selection procedure to identify the best model, or just use the climate trend (from the climate Bayesian DFA) as the climate covariate as it represents the shared variability of all the climate variable.

Response: Thank you for this comment. We are planning to update our analysis annually to provide an indicator of ecosystem state for the California Current Integrated Ecosystem Assessment and Ecosystem Status Report. We intend to expand our modeling framework and conduct a variable selection procedure as you described. We mention this and other extensions of our study in the discussion section starting at line 853. We prefer the variable selection approach using the ROMS output over using the climate trend because using ROMS outputs will enable us to use ROMS forecasts to then forecast biological changes in the CCE.

13. Line 353. It should be S2 Fig.

Response: Thank you for catching this. We made this change.

14. Lines 383-386. According to the “90%” criterion, Pacific sardine is not or weakly associated with the trend. Therefore, there is less need to focus on these species that were not or weakly captured by the trend as their dynamics are not well reflected by the trend.

Response: All species shown in figure 4b loaded strongly on the trend (probability > 0.9), including juv/adult Pacific sardine (see lines 450-459). Species that were included in our analysis but did not load strongly on this trend (probability < 0.9) are not shown in figure 4b and are not mentioned in the results or the discussion.

15. Lines 395-398. Similar to question 7, as different span of time series will cause uncertainty, why do not the authors unify spans of all the time series.

Response: We used the full length of the available time series for the biology DFA model because we wanted to produce a historical record, dating as far back as possible, of the community state in the California Current Ecosystem. Despite the large uncertainty around the trend earlier in the period (e.g. 1950-1970), the trend values during this time are still useful for evaluating changes in community state over time and evaluating whether they are within the range of normal variability or not. We note that for the biology-climate covariate models, we unified the time spans of the climate and biology time series.

16. Lines 405-409. I wonder if there is any similarity in these species as they show synchronous increases.

Response: Good question. Three of these species are cool water associated mesopelagic fishes, which were sampled in the southern California Current. We added this information to lines 478-482.

*This shift coincides with a strong increase in the abundance of a few species during that period, including eared blacksmelt (*Lipolagus ochotensis*), slender blacksmelt (*Bathylagus pacificus*), northern lampfish (*Stenobrachius leucopsarus*), which are cool water associated mesopelagic species, as well as a rise in northern anchovy (*Engraulis mordax*) abundance prior to the shift (S1 Fig.).*

17. Line 412. It should be S3 Fig.

Response: Thank you for catching this. Done.

18. Line 429. The term B matrix is better to appear in the materials and methods.

Response: We removed the B matrix term from this sentence (line xx)

19. Line 437 It should be S4 Fig.

Response: Thank you for catching this. Done.

20. Lines 436-437. The result of biology-CUTI model is similar to the biology-BEUTI model, it may indicate the two covariate is correlated. Therefore, a correlation analysis among climate variables can help readers understand your results better.

Response: Yes, this is correct. These variables are correlated, and we mention this in the discussion on lines 853-858.

21. Line 564. The “Suryan et al., 2021” is not found in the reference list.

Response: Thank you for catching this. We added this citation to the reference list.

22. Lines 710-731. The discussion about how we can use the results from this study to benefit management in the CCE is highly recommended.

Response: The final three paragraphs of our discussion are devoted to explaining how our study is applicable to ecosystem-based management in the CCE (lines 959-1024). We added a new heading 'Management application' at line 959 and new text at lines 965-967.

Our intention is to continually update our analyses when new data become available to provide the most up-to-date information on the CCE community state for scientists, managers, and stakeholders.

We feel that these paragraphs are sufficient for linking our work to management, however, we will provide additional text if the Editor thinks it's necessary.

Reviewer #2: Comments on “Tracking and forecasting community responses to climate perturbations in the California Current Ecosystem”

General comments:

The manuscript implemented a Bayesian Dynamic Factor Analysis method to track the response of California Current Ecosystem to climate perturbations and to forecast future changes in community state. Their results demonstrated strong relationships between community state and multiple climate variables, and identified that nitrate flux through the base of the mixed layer had the strongest correspondence with individual species and the community trend.

Overall, I think the paper is well written and logically presented. The analyses are conducted with solid statistical methods and supporting data. My concerns lie on the organization of the results and discussions, all of which are connected to the objectives of this study. The study included a rich results, however, a large of texts usually make the readers difficult to follow the center idea of the paper. The author may consider to focus on the most impressive results to make it concise, by removing the less ones. This is a following result of the unclarified objectives in the introduction, where the authors attempt to include a range of goals in this study.

*Please note: All line numbers in our responses below refer to the revised manuscript with tracked changes.

Response: Thank you for your constructive comments. We are aware that there is a lot of information presented in this manuscript. We have clarified our goals and objectives in the final paragraph of the introduction. We also present the reader with a road map of our work flow in the *Modeling* sub-section of the methods section and we indicate that the four steps of our work flow map on to the four study objectives outlined in the introduction (see lines 244-252).

The discussion is tedious. I appreciate the authors' efforts to organize the sweeping discussion, where as it should be pointed out that no all of them are necessary. I would suggest the authors to focus on the most important findings in the discussion, and if possible, try to bring new insight by summering the underlying patter and driving factors. A list of many relevant facts makes little help towards in-depth discussions.

Response: Thank you for your suggestions. We trimmed parts of our discussion based on the reviewer's comments. We also introduced section headers within the discussion to help the reader follow the organization and break up the paragraphs. We note, however, that we were less inclined to remove/re-organize text pertaining to the forecasting component of the paper. The forecasting component of the DFA is a novel development of our modeling approach and therefore we think it's appropriate to include most of the existing text. It may seem tedious to

some but we think it is valuable for readers who are interested in applying a similar approach to their own study systems.

When novel community states were considered, I wonder if it would be an issue that the very strong signals in the past may conceal any signs of changes in recent years, resulting in insignificant results, which however does not necessarily mean the lack of alternate community state. The point here is the statistical significance may influence each other if the extent of changes differ substantially among the community states, which is dependent on the variables measured in the surveys.

Response: This is a really great point, and something we're actively working on to improve with these models. The type of model you're describing is one with non-stationary variability -- e.g. the magnitude of the variation is changing in some way, either randomly, or over time. If the non-stationary process is estimable, then output from that model would allow the meaning of each state to also change over time (so that strong signals in the past wouldn't necessarily impact inference of change in recent years). We discuss the issue of non-stationarity in our discussion on lines 953-957).

Specific comments:

1. Line 42-49: the objectives may be summarized or stated in brief in the introduction.

Response: The specific objectives of the study are listed in the final paragraph of the introduction.

2. Line 77: The texts after "Since then" may be organized in the following paragraph.

Response: Done.

3. Line 95: it seems arbitrary to attribute all the events mentioned to the consequence of heatwaves.

Response: We respectfully disagree. We feel that it's appropriate to acknowledge the previous studies and findings that motivated this study. While it may be possible that the conclusions of some of these studies are uncertain, the consistency among them with respect to the ecosystem-wide consequences of this event are robust. Readers are, of course, free to read the original reports and arrive at their own conclusions.

4. Line 105: it is not quite clear how the early detection is viable here.

Response: The idea here is to provide the earliest possible detection that the response of the broader ecological community to a climate perturbation is outside the observed range of normal variability and therefore may be a signal of a wider ecosystem shift and/or an indication to increase ecosystem monitoring and take a more precautionary approach to management.

5. Line 110: No doubt there are promising progress in community modelling. Even so, I can hardly agree that “the time is ripe” for near-term forecasts, regarding the uncertainty in current modelling frameworks and the limited knowledge in community dynamics.

Response: We changed the text (line 115) to “*as climate models and forecasts of ocean conditions continue to improve, there are burgeoning opportunities to develop and test methods that could provide near-term forecasts of community state...*”

6. Line 114: “asynchrony among time series” may be explained in more details.

Response: Here we mean that the time series data are unevenly or irregularly spaced.

7. Line 117: PCA and MDS represent a small partition of the method used in the research field of community ecology, and a more comprehensive review is recommended on this topic.

Response: This comment refers to “tools such as Principal Components Analysis (PCA) or nonmetric multidimensional scaling have often been used for identifying leading patterns of variability in multivariate datasets (e.g., Koslow et al. 2002, 2013)”. Our text acknowledges that these are merely examples. A review of the different methods and the strengths/weaknesses of each is beyond the scope of this paper. Zurr et al. 2003, which we cite in this paragraph, also addressed this point so it’s not new. Our point here is simply that DFA is better suited for time series analysis than some of the multivariate approaches that have been commonly applied to marine ecosystems. We provided additional information on why it’s a better method in response to reviewer 1 comment above (#4):

Ward et al. (2019) recently developed a Bayesian implementation of DFA that offers added flexibility in model aspects over conventional approaches; examples include allowing for extreme “black swan” events (rare and difficult to predict events; Anderson et al. 2017), and trend processes that don’t follow a random walk. Output from these Bayesian DFA models can also be used to estimate the probability of extreme events occurring or switches among contrasting system states.

8. Line 140: the goals may be revised and reorganized to be concise, distinguishing between approaches and objectives.

Response: We changed some of the text in the final paragraph of the introduction to better distinguish the goals of the paper from the specific objectives of our study (lines xx-xx)

9. Line 169: In this case, it would be critical how the ROMS was calibrated and validated in the ecosystem. Some supporting information are recommended.

Response: We inserted the following text at lines 199-203:

This ocean model is constrained by available satellite and in situ observations to improve its fidelity to nature, and has been validated against independent in situ observations (Neveu et al. 2016, Schroeder et al. 2014). Output from this model has been widely used to characterize CCS oceanography, its relation to large scale climate variability, and its influence over the marine ecosystem from phytoplankton to top predators (see Discussion).

10. Line 187: how is the threshold of 15 years selected?

Response: For our analysis, we wanted to include long time series that spanned as many climate perturbations as possible and also have enough biological time series to develop an informative indicator of community state. The threshold of 15 years allowed us to meet both of those criteria. In addition, 15 years is a threshold that has been previously used to define "long oceanographic time series" in the California Current. For example, McClatchie et al. 2014 wrote "Here, we define long time series in fisheries oceanography as surveys longer than 15 years that sample either juvenile fish or ichthyoplankton (fish eggs and larvae) combined with fields of oceanographic properties."

We added the following text to lines 221-226:

A threshold of 15 years allowed us to include long time series that spanned as many climate perturbations as possible and also have enough biological time series to develop an informative indicator of community state. In addition, 15 years is a threshold that has been previously used to define "long oceanographic time series" in the California Current (McClatchie et al. 2014).

McClatchie, S., Duffy-Anderson, J., Field, J.C., Goericke, R., Griffith, D., Hanisko, D.S., Hare, J.A., Lyczkowski-Shultz, J., Peterson, W.T., Watson, W. and Weber, E.D., 2014. Long time series in US fisheries oceanography. Oceanography, 27(4), pp.48-67.

11. Line 187: It should be noted that index standardization is usually used for fishery-dependent surveys in which data are usually aggregated in resource-rich areas, but not for properly designed scientific surveys.

Response: There is some interannual variability in the sampling effort (time and space) of the scientific surveys associated with our study and so it was appropriate to standardize the data for those surveys. We also would highlight that index standardization is routinely used for fishery-independent surveys around the world — these surveys are ‘properly designed’ but the samples are spatially random year to year, and not at specific stations. For example, see the following papers:

Maunder, M.N. and Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. *Fisheries research*, 70(2-3), pp.141-159.

Potts, S.E. and Rose, K.A., 2018. Evaluation of GLM and GAM for estimating population indices from fishery independent surveys. *Fisheries Research*, 208, pp.167-178.

Thorson, J.T., Maunder, M.N. and Punt, E., 2020. The development of spatio-temporal models of fishery catch-per-unit-effort data to derive indices of relative abundance. *Fisheries Research* 23: 105611

12. Line 197: it seems questionable to change all zeros to NAs, as zeros may be observations rather than missing values in surveys.

Response: We had considered other approaches, such as adding a very small number to the response before log-transforming. However, this is known to cause issues and be sensitive to the choice of arbitrary constant. By changing the values to NAs, we’re assuming that the species was within the survey region and available to be detected, but was in such small quantities that it was not. The choice of approach is a question of professional judgement, but one that the results are generally robust to.

13. Line 199: the influences of log transformation on the DFA may be noted here.

Response: The log-transformation is common in fisheries applications of DFA, because the objective is to create time series that have approximately normal observation and process errors. We have included additional distributional families in the R package ‘bayesdfa’, but have not yet tested these — future modeling work will ideally include these to work with raw distributions that vary by time series.

14. Line 234: could nonlinear relationships be handled in the modelling framework?

Response: Yes, non-linear modelling relationships are simple to include. We have some examples demonstrating this in the ‘bayesdfa’ package — these can be done with smooth splines, like a GAM, or Gaussian processes, etc.

15. Line 249: should the climate datasets and biological datasets be used in the same time span?

Response: The time spans don’t need to be the same to estimate separate biology and climate trends. We wanted to create a historical record of the community state dating as far back as possible. However, we did truncate the biology time series to match the time span of the climate times for the biology-climate covariate models.

16. Line 259: more details may be needed about the LFO-CV before explaining the target of the method.

Response: We think this is clear as written — additional technical details can be found in the cited Burkner et al. paper.

17. Line 270: may denote it as T-1. Also, could the prediction be considered for a scope more than one year?

Response: We changed the text at line 313 to: *For each model formulation, we applied the LFO-CV method by first fitting the model to all years of data prior to year T (i.e., training data, years 1, 2, ..., (T-1)) and then using the fitted model to predict the trend value in year T (i.e., test data).*

We could extend the modeling approach to make predictions more than one year in advance; however, we chose to start with one year as that would provide the best test for identifying whether we have any forecast skill. If we don’t have much skill forecasting the biological response one year ahead then we’re not likely to have much skill forecasting the response a few years in advance. Also, one-year-ahead forecasts are appropriate for timing of management decision making in the CCE, which happens on an annual basis.

18. Line 271: “repeated this process for a time series dataset of 10 years” or something like this may be clearer.

Response: We changed to ‘repeated this process for 10 years of the dataset’.

19. Line 274-277: show them in an earlier contexts. A subtitle of “model structure optimization” may be added to this section somewhere.

Response: We added the subtitle ‘Model structure optimization’ to line 290.

20. Line 291: How does this LOOIC relate to the LFO-CV mentioned earlier?

Response: LFO-CV was used for model selection of the climate, biology, and biology-covariate DFA models because these models are intended to be used in a forecasting context. LOOIC should have been written as LOO-CV (sorry for any confusion). LOO-CV is the Leave-One-Out Cross Validation information criterion used for model selection of the Hidden Markov Models that we used to estimate the presence of regimes in the state indices.

21. Line 298: could have more texts about the “community states and raw time series” here.

Response: We clarified our meaning by adding text to lines 339-342.

While a wide variety of multivariate or univariate time series methods could be applied to our observed time series to generate forecasts, our objectives were to develop simultaneous estimates of both the community state (i.e., the biology trend value) and the raw time series (i.e., individual time series summarized by the biology DFA model).

22. Line 311: the “individual species parameters and the community state” are not clear here.

The last sentence of this paragraph may be revised.

Response: We clarified our meaning with the following text (lines 353-371):

Once the best-supported biology-covariate model was identified, we used that model to make predictions of the community state (i.e., DFA trend value) in 2018 using climate data from that same year and the raw time series of the individual species (i.e., the biology time series summarized by the DFA model). We evaluated forecast skill based on the prediction errors of individual species time series and by comparing the forecasts for 2008–2018 to the 2008–2018 trend values estimated from the biology-covariate model that only included data prior to the forecast year.

23. Line 327: A description of trend may be placed before the correlation with climate time series.

Response: We switched the order of the text here (lines 375-389).

24. Line 359: not quite clear how the number of states were determined here.

Response: The number of states was determined based on the Hidden Markov model with the lowest LOO-CV value. We now mention this in the method on line 334-336.

25. Line 382: may mark the occurrence of the marine heatwave in the figure.

Response: We made a note of the marine heatwave years in the figure 4 caption.

26. Line 386: supplementary figures may be needed to illustrate the associations with the heatwave, as Figure 4 only showed the correlation with the single trend.

Response: We're unsure about what information the reviewer is suggesting we present in supplementary figures. This text provides an ecological interpretation of the loadings with respect to the biology trend. The discussion provides additional context for our interpretation of biological / community response to the marine heatwave.

27. Line 397-401: some of them can be moved to the discussion or SI, regarding the rich texts in the results section.

Response: We removed this text (original line numbers 398-401) because this information is also presented in the discussion. Otherwise, the text is directly relevant to the main findings of the paper.

28. Line 424-473: consider cut short the texts in this section. Some results may be shown in the SI if they are not closely related the topic of this study.

Response: The text in this paragraph is directly relevant to the topic of this study. We moved Fig. S8 (now Figure 7) to the main text to emphasize that these results are a key finding of our paper. We note that an older version of Fig. S8 was included in the original version of this manuscript by mistake. We now include the correct version of this figure, which is slightly different but does affect the results in any way.

29. Line 480: explanation of prediction uncertainty may be placed in the discussion.

Response: Given that the DFA model forecasts are a novel component of our modeling framework, we think it's important to include this explanation within the results section.

30. Line 508-548: the discussion may focus on the potential insights that the study can provide, instead of listing all the relevant facts per se. otherwise, the discussion can be largely condensed.

Response: We trimmed and re-organized this paragraph based on the reviewer's comments.

31. Line 558: “Suryan et al. (2021)” was not properly included in the references.

Response: Thank you for catching this. We have included this citation in the references section.

32. Line 596: the discussions could be more focused on the interpretation of identified DFA trends in the biological and ecological sense.

Response: Under the new subheading ‘Long-term changes in community state’, we provide a good interpretation of the ecology and potential mechanisms driving the trends and loadings and covariate effects. Also, because our study is an attempt to synthesize a broad suite of community indicators and their response to climate perturbations our intent is not to get too detailed in our interpretation of each indicator included here. Instead, we point readers to more detailed investigations into the drivers and mechanistic understanding of the indicators.

33. Line 650-669: this paragraph may be removed as it is not exactly relevant to the analyses of this study.

Response: We respectfully disagree with this comment. An important goal of this paper is to test our ability to develop near-time forecasts of community state using CCE ROMS output and forecasts. This paragraph provides support for using CCE ROMS to forecast biological processes in the California Current. This paragraph also addresses this reviewer’s comment #9.

34. Line 696-709: This is also not quite a proper discussion for this study, which is not designed for the purpose of methodological development

Response: Again, we respectfully disagree as this paragraph provides further information on the forecasting approach and how it might be modified to address different but related questions. Moreover, the forecasting component of the DFA is a novel methodological development.

Journal Requirements:

1. Please provide separate figure files in .tif or .eps format only, and remove any figures embedded in your manuscript file. If you are using LaTeX, you do not need to remove embedded figures.

Response: We converted all figures for the main text from .pdf files or .tif files. The figures are also available as .eps files, if needed,

2. In the online submission form you indicate that your data is not available for proprietary reasons and have provided a contact point for accessing this data. Please note that your current contact point is a co-author on this manuscript. According to our Data Policy, the contact point

must not be an author on the manuscript and must be a third party. Please revise your data statement to a non-author institutional point of contact, such as a data access or ethics committee, and send this to us via return email. Please also include contact information for the third party organization, and please include the full citation of where the data can be found.

Response: We did not include authors as contact points to provide expert knowledge on the surveys and biological collections should readers have any questions. We removed this information and instead only include direct links to the databases/dashboards where data can be accessed and/or a non-author institutional point of contact for accessing data.

3. Please provide us with a direct link to the base layer of the map used in Figure 1 and ensure this location is also included in the figure legend.

Response: We updated the map in Figure 1 using a base layer from NOAA that is in the public domain. We provide the following text to the Figure 1 caption.

The base map layer was sourced from [NOAA National Geophysical Data Center \(2009\) ETOPO1 1 Arc-Minute Global Relief Model](#). NOAA National Centers for Environmental Information (accessed: 19 April 2013, Amante, C & BW Eakins 2009).

Amante, C & BW Eakins (2009) ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NESDIS NGDC-24, National Geophysical Data Center, United States Department of Commerce, Boulder, CO, March 2009. 19 p. Accessed: 19 April 2013. DOI: [10.7289/V5C8276M](https://doi.org/10.7289/V5C8276M)

4. Please amend your detailed Financial Disclosure statement. This is published with the article, therefore should be completed in full sentences and contain the exact wording you wish to be published.

i). State the initials, alongside each funding source, of each author to receive each grant.

ii). State what role the funders took in the study. If the funders had no role in your study, please state: “The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.”

Response: We amended the statement to read:

Funding for this project came from NOAA’s Fisheries and the Environment (FATE) program (project 16-01) awarded to M.E.H, E.J.W., M.A.L. and C.J.H. and NOAA’s California Current Integrated Ecosystem Assessment program (C.J.H.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

5. Changes to the reference list should be mentioned in the rebuttal letter.

We added the following new citations to our reference list.

1. Amante C, Eakins BW.ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NESDIS NGDC-24, National Geophysical Data Center, United States Department of Commerce, Boulder, CO, March 2009. 19 p. DOI: 10.7289/V5C8276M
2. Checkley DM Jr, Barth JA. 2009. Patterns and processes in the California Current System. *Prog. Ocean.* 2009; 83, 49-64.
3. Gottscho AD. Zoogeography of the San Andreas Fault system: Great Pacific Fracture Zones correspond with spatially concordant phylogeographic boundaries in western North America. *Biological Reviews* 2016; 91: 235-254.
4. Jacox MG, Fletcher J, Moore AM, Edwards CA. ENSO and the California Current coastal upwelling response. *J. Geophys. Res.* 2015. doi:10.1002/2014JC010650
5. McClatchie S, Duffy-Anderson J, Field JC, Goericke R, Griffith D, Hanisko DS, Hare JA, Lyczkowski-Shultz J, Peterson WT, Watson W, Weber ED. Long time series in US fisheries oceanography. *Oceanography* 2014; 27: 48-67.